Interactive Model-Centric Systems Engineering (IMCSE)

Phase 4

Technical Report SERC-2017-TR-103

March 1, 2017

Principal Investigator:

Dr. Donna H. Rhodes, Massachusetts Institute of Technology

Research Team:

Dr. Adam M. Ross, Massachusetts Institute of Technology

Michael Curry, Massachusetts Institute of Technology

Matthew Fitzgerald, Massachusetts Institute of Technology

Erling Shane German, Massachusetts Institute of Technology

Jack Reid, Massachusetts Institute of Technology

Lucie Reymondet, Massachusetts Institute of Technology

Sponsor: DASD (SE)

The Systems Engineering Research Center (SERC) is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Office of the Assistant Secretary of Defense for Research and Engineering (ASD(R&E)) under Contract HQ0034-13-D-0004 (TO 0062).

TABLE OF CONTENTS

Table of Contents iii

List of Figures v

List of (Tables, Sequences) vii

Executive Summary 1

Interactive Model-Centric Systems Engineering (IMCSE) Research 3
- **Pathfinder** 3
- **Phase 4 Research Focus** 4

Interactive Epoch-Era Analysis 5
- **Background** 5
- **IEEA framework** 7
- **Case Application: Commercial Offshore Ship** 8
  - Process 1: Value-Driving Context Definition 9
  - Process 2: Value-Driven Design Formulation 9
  - Process 3: Epoch Characterization 10
  - Process 4: Era Construction 10
  - Process 5: Design-Epoch-Era Evaluation 10
  - Process 6: Single Epoch Analysis 11
  - Process 7: Multi-Epoch Analysis 12
  - Process 8 and 9: Single and Multi-Era Analysis 14
- **Impact Assessment: Human-in-the-Loop Software Study** 15
  - Treatment Groups 16
  - Trial Protocol 19
  - Experimental Task 20
  - Interim Results 20
- **IEEA: Discussion and Future Research** 21

Model-Centric Decision Making Study 22
- **Motivation** 22
- **Interview-Based Study Approach** 23
- **Decision-Making Flow** 25
- **Trust in Models** 26
- **Interim Findings** 27
  - Technological and social factors influencing trust 28
  - Importance of communication 28
  - Transparency 29
  - Understanding of assumptions and uncertainty 29
  - Effective Model Documentation 30
  - Primary versus Supplementary 30
  - Independent review 31
  - Investment bias 31
  - Confirmation bias 31
  - Endogeneity of the human 32
  - Real-time interaction with models 32

Preliminary Heuristics (Guiding Principles) **33**
Using Models in Decision Making 33
Importance of Model Context and Assumptions 33
Modeling Process and Decisions 34
Transparency and Trust 34
Model-Centric Environments 34
Model-Centric Decision Making Study: Discussion and Future Research **35**

Recommendations for Framing Multi-Stakeholder Tradespace Exploration **36**
Introduction **36**
Multi-Stakeholder Tradespace Exploration **37**
Macro Framing and Micro Framing **38**
Framing Activities and Visualizations **42**
Problem Formulation 42
Modeling / Evaluation 44
Exploration / Analysis 45
Framing Multi-Stakeholder TSE: Discussion and Conclusion **50**

Model Curation **52**
Background **52**
Driving Factors **55**
Knowledge Areas **56**
Model Pedigree **57**
Model Curation Role **58**
Lexicon (preliminary) **59**
Model Curation: Discussion and Future Research **60**

Conclusion **61**

Appendix A: List of Publications Resulted **63**

Appendix B: Cited and Related References **64**

LIST OF FIGURES

Figure 1 IMCSE Motivations and Foundations ... 1

Figure 2 IMCSE Pathfinder Workshop Has Informed Research Areas ... 3

Figure 3 Areas of Inquiry ... 4

Figure 4 Framework Models and Processes ... 7

Figure 5 System boundaries and epoch variables (Rehn et al., 2016) ... 10

Figure 6 : Interactive Filtering Application for tradespace exploration for the offshore ship design base case ... 11

Figure 7 Interactive heatmap visualization (left), inspection table (right) for ships in selected tile in heatmap. ... 13

Figure 8 Interactive Filtering Application: implementing OLAP for the offshore case (Filtered for allowable change cost < $100M and change time < 90 days) ... 14

Figure 9 Interactive Multi-Era Application implementing OLAP for the offshore case ... 15

Figure 10 Experimental Treatment Groups ... 17

Figure 11 Treatment A - Non-Interactive Table ... 17

Figure 12 Treatment B - Non interactive Table + Visualization ... 18

Figure 13 Treatment C - Interactive Table ... 18

Figure 14 Treatment D – Interactive Table + Visualization ... 19

Figure 15 Trial Protocol ... 20

Figure 16 Elements of Model-Centric Decision Making ... 22

Figure 17 IMCSE Exploratory Study on Model-Centric Decision Making ... 23

Figure 18 Three actor decision flow ... 26

Figure 19 Technological and Social Factors Influencing Trust ... 28

Figure 20 Perceived value around a reference point, according to Prospect Theory ... 40

Figure 21 Negotiation tradespace used in MSTSE experiment (Fitzgerald and Ross, 2015) with key features ... 46

Figure 22 Example Venn diagram and Gridmap for multi-stakeholder Pareto front analysis ... 47

Figure 23 Example Stakeholder-Stakeholder correlation interface for five stakeholders – annotated to highlight two emergent coalitions with high correlation (orange and magenta boxes) ... 48

Figure 24 Example Stakeholder-Interest correlation interface – annotated to highlight the attribute most responsible for bringing the orange coalition together and separating them from magenta ... 49

Figure 25 Example interface for live editing of value functions, synced to other visualizations . 50
Figure 26 Many Decisions in Using Model-Centric Environments 52
Figure 27 Enterprise-level curation of model-centric environments 55
Figure 28 Knowledge Areas for Model Curation 56
Figure 29 IMCSE Research Assumptions.. 61

LIST OF (TABLES, SEQUENCES)

Table 1 List of Interview Questions 24

Table 2 Summary of recommendations, with modifications for informal MSTSE 51

Table 3 Model Curation Lexicon (selected examples) 59

EXECUTIVE SUMMARY

The Interactive Model-Centric Systems Engineering (IMCSE) research program arises from the unique opportunity to investigate the various aspects of humans interacting with models and model-generated data, in the context of systems engineering practice. IMCSE research aims to develop transformative results through enabling intense human-model interaction, to rapidly conceive of systems and interact with models in order to make rapid trades to decide on what is most effective given present knowledge and future uncertainties, as well as what is practical given available resources and constraints. While model-based engineering initiatives are advancing technical aspects of models in the engineering of systems, this research advances knowledge relevant to human interaction with models and model-generated information.

As portrayed in Figure 1, increasing use of models motivates the investigation of interactive model-centric systems engineering. IMCSE research pulls foundational knowledge from other fields such as model-based systems engineering, complex systems, big data science, and visual analytics.

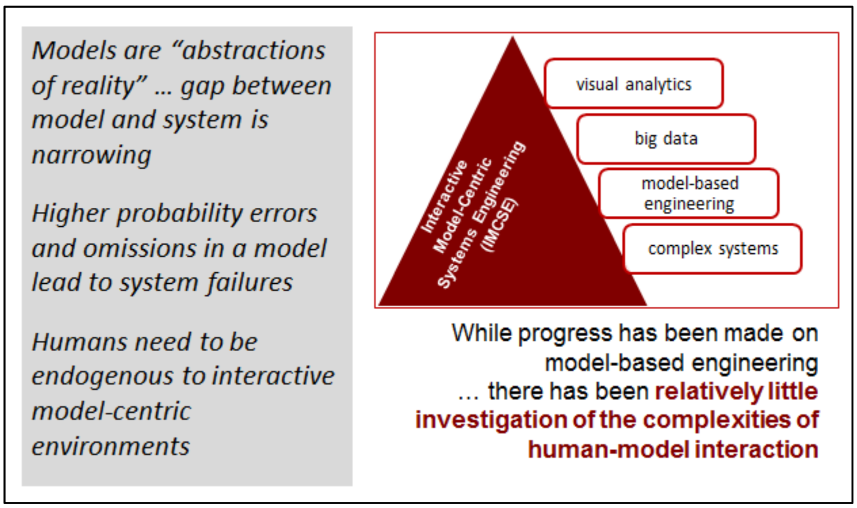


Figure 1 IMCSE Motivations and Foundations

This report discusses results of Phase 4 of the ongoing IMCSE research program, focusing on four areas:

Interactive Epoch-Era Analysis. Continued research was performed on an innovative method for evaluating systems under dynamic uncertainty using epoch-era analysis with focus on enhanced interactive capability and allowing for scaling for big data analysis. The Interactive Epoch-Era Analysis framework and supporting tools were applied to a commercial offshore ship, demonstrating key concepts and prototype interactive visualizations, focusing on opportunities to improve the uncertainty analysis, ease of use, data scaling, visualization techniques, and overall analysis approach.

Model-Centric Decision Making. A study was initiated to generate empirical insight into how various human actors, including decision-makers trust, perceive, and interact with models. An interview-based approach is used to identify important considerations surrounding human-model interaction and trust that experts deem important for effective decision-making. These considerations include practices that interviewed experts use to aid in their decision-making, along with identified challenges that can degrade effective model-centric decision-making. The descriptive insights gained through empirical research, along with research on decision-making and biases, aims to identify heuristics and guiding principles to model-centric engineering policy and practice.

Framing Multi-Stakeholder Tradespace Exploration. Tradespace exploration (TSE) and, specifically, multi-stakeholder tradespace exploration (MSTSE) support early conceptual design of engineering systems with multiple stakeholders. MSTSE has been developed to target design tasks with stakeholders who are unwilling or unable to fit their preferences into a shared normative decision framework but who remain involved in the design process. Framing has been identified as a challenge leading to counterproductive negotiation tactics by previous MSTSE research, but a challenge that is capable of being ameliorated through creative redirection of attention and emphasis on group-dynamic data over individualistic data. Recent research has resulted in recommendations for framing adjustments throughout the MSTSE process, including early in the problem formulation.

Curation of Model-Centric Environments. As model-centric environments become increasingly complex and important, there is a need to more strategically lead and manage them. Under the premise that model-centric environments of the future will necessitate specialized leadership and competencies, a new leadership role for curation has been investigated. The curation function would set and administer model-related policies and practices, and ensure models and related documents are authenticated, preserved, classified and organized. The curator may own the data management for models and related information, or oversee the ownership by other individuals or organization. As needed, a curator would meet with individuals and teams, who will create, use and re-use digital assets, helping to determine a useful classification of both individual models and sets of models. At the organization level, the curator may organize training and special projects. Empirical knowledge gathering has investigated the challenges and needs, and investigated the potential roles and responsibilities for this curation function.

During this phase, IMCSE research was presented and discussed with practitioners and sponsors in numerous research meetings and workshops. A SERC Talk, *Why is Human-Model Interactivity Important to the Future of Model-Centric Systems Engineering?*, highlighted various research efforts under the project. Continued research and knowledge transfer is raising the awareness of challenges and needs surrounding human-model interactivity, and there is a growing community of interest with the SERC and the larger systems community.

INTERACTIVE MODEL-CENTRIC SYSTEMS ENGINEERING (IMCSE) RESEARCH

The Interactive Model-Centric Systems Engineering (IMCSE) research program, initiated in 2014, aims to inform and contribute methods, processes and tools to improve human-model interaction, in support of accelerating the transition of systems engineering to a more model-centric discipline (Rhodes and Ross, 2015). The IMCSE research program arises from the opportunity to investigate the various aspects of humans interacting with models and model-generated data. IMCSE aims to develop transformative results through enabling intense human-model interaction, to rapidly conceive of systems and interact with models in order to make rapid trades to decide on what is most effective given present knowledge and future uncertainties, as well as what is practical given resources and constraints. Additionally, this research generates knowledge impacting human effectiveness in model-centric environments of the future (Rhodes and Ross, 2016). Future environments and practices need to leverage advancements in data science, visual analytics, and complex systems.

The transformation of systems engineering to a model-centric paradigm is progressing at a rapid pace. Models are increasingly used to drive major acquisition and design decisions, yet the diverse set of model developers, analysts, and decision makers are faced with many challenges. The systems community has made progress on standards, methods and techniques for model-based systems engineering, yet little focus has been given to complexities of human-model interaction. A science of human-systems integration (HSI) has emerged (Pew and Mavor, 2007), yet focus is on humans within operational systems, while models are abstractions of reality. The relatively mature field of human-computer interaction (HCI) offers valuable insights (Harper et al., 2008), however focus is on designing computer interfaces.

PATHFINDER

The IMCSE Pathfinder Workshop held in January 2015 gathered research stakeholders for an initial dialogue on the topic (Rhodes and Ross, 2015). During the event, the attendees considered an envisioned future for model-centric engineering. A number of emergent themes from the workshop discussions have informed the directions of the research program. Figure 2 enumerates some of these themes.

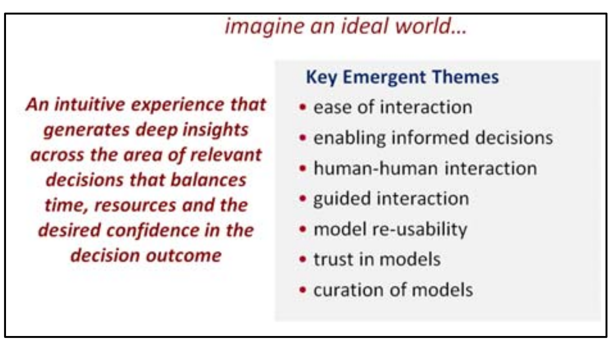


Figure 2 IMCSE Pathfinder Workshop Has Informed Research Areas

Open areas of inquiry include: how individuals interact with models; how multiple stakeholders interact using models and model generated information; facets of human interaction with visualizations and large data sets; and underlying fundamentals such as model purpose and model handling. IMCSE research is based on a belief that human-model interaction needs to be specifically considered, given models are an abstraction of reality and there are likely unique factors and considerations. Realizing the envisioned future for an interactive model-centric systems engineering experience will require new knowledge and ways of working, and innovation in constructs and technologies. Figure 3 lists several questions the research seeks to address.

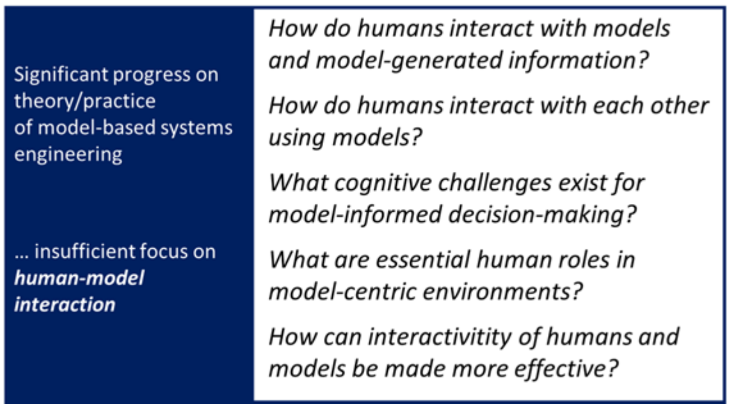


Figure 3 Areas of Inquiry

PHASE 4 RESEARCH FOCUS

The Phase 4 research provided the opportunity to build on the research outcomes for the prior phases. This report discusses the findings of four areas of focus:

1. Interactive Epoch-Era Analysis
2. Model-Centric Decision Making
3. Recommendations for Framing Multi-Stakeholder Tradespace Exploration
4. Model Curation

INTERACTIVE EPOCH-ERA ANALYSIS

Epoch-Era Analysis (EEA) is a framework that supports narrative and computational scenario planning and lifecycle uncertainty analysis for both short run and long run futures. Because of the complex data that must be analyzed when extending EEA to large-scale problems, issues with cognition are introduced that may hamper decision-making. This motivates the need for extensions to EEA methods that overcome the computational and human cognition issues that may arise. The Interactive Epoch-Era Analysis (IEEA) framework, comprised of 10 processes grouped in 6 modules, is introduced as a means for analyzing lifecycle uncertainty when designing systems for sustained value delivery. IEEA is proposed as an iterative framework for concept exploration that provides a means of applying EEA constructs while controlling growth in data scale and dimensionality. Further, IEEA leverages interactive visualization because prior visual analytics research has demonstrated that when performing exploratory analysis, like early-phase system concept selection, an analyst can gain deeper understanding of data which can lead to improved decision-making. In the prior phase the framework and supporting tools were applied to a multi-mission on-orbit servicing vehicle (Curry and Ross, 2016), demonstrating key concepts and prototype interactive visualizations, focusing on opportunities to improve the uncertainty analysis, ease of use, data scaling, visualization techniques, and overall analysis approach. During this phase of the research it was applied to a commercial ship case to further test the framework and refine visualization prototypes. An impact assessment experiment was initiated to decouple and evaluate the impacts of visualization and interaction on human performance

BACKGROUND

Epoch-Era Analysis (EEA) is designed to clarify the effects of changing contexts over time on the perceived value of a system in a structured way (Ross and Rhodes, 2008). The base unit of time in EEA is the epoch, which is defined as a time period of fixed needs and context in which the system exists. This approach provides an intuitive basis upon which to perform analysis of value delivery over time for systems under the effects of changing circumstances and operating conditions, an important step to take when evaluating large-scale engineering systems with long lifecycles. Interactive Epoch-Era Analysis (iEEA) leverages human-in-the-loop (HIL) interaction to manage challenges associated with the large amounts of data potentially generated in a study, as well as to improve sense-making of the results (Curry and Ross, 2015). Allowing the structured evaluation and visualization of many design alternatives across many different futures and potential lifecycle paths enables the design of systems that can deliver sustained value under uncertainty. Iteration is necessary because the analysis is inherently exploratory in nature. HIL interaction is necessary because the problem is not strictly deterministic or necessarily intended as a reliable prediction of system performance or future events. Often, there is both uncertainty and the potential for errors in assumptions or model implementation. This necessitates human judgment to make sense of the data; therefore, this is not by its nature a problem that can be handed over completely to an automated optimization algorithm, though some level of automated analysis could be beneficial as an aid

to the user. Enabling users to interact with their data through visual interfaces of this type is an area of active research (Heer and Agrawala, 2007, Heer and Shneiderman, 2012). Problems with rendering and the scalability of visualizations and other encoded visual information can be improved upon using techniques that do not require every single data point to be drawn. Liu points out that, “Perceptual and interactive scalability should be limited by the chosen resolution of the visualized data, not the number of records,” and summarizes several techniques past researchers have applied to reduce the pixel density of visualizations including (1) filtering; (2) sampling; (3) binned aggregation; and (4) model-fitting (Liu et al., 2013). Online Analytical Processing (OLAP) is an approach for creating abstract representations of high-dimensional datasets. OLAP is frequently applied in data mining and other exploratory analysis applications with large amounts of data. These datasets are often stored in relational databases with multiple tables connected by keys, but can also be as simple as a spreadsheet with records stored in each row and with columns representing different attributes or properties of the data. In fact, pivot tables generated in MS Excel are one example of a common application of OLAP for summarizing data. A notable application of OLAP is its successful use in business intelligence applications to parse large amounts of sales, cost and other data to evaluate trends and inform business decisions. For IEEA, the benefit of OLAP is that it enables a user to view data from multiple points of view and quickly uncover previously undiscovered relationships and patterns within the dataset. A decision-maker looking at a large number of candidate designs across a large possible epoch space can apply OLAP techniques to slice, dice, drill down, roll up or compute pivots of the hyper-dimensional data cube representing design alternatives over epochs and eras. This allows them to easily extract data that is of interest to them which, in turn, enables better intuition on which to base decisions. Multiple coordinated views can be used in exploratory visualization to more effectively expose relationships in the underlying data. Coordinated views are separate, independent views of a given set of data that serve as complementary representations, and may aid in identifying patterns as well as errors in the data. The individual views of the data are not intended for use in isolation, but rather to be combined to generate insights. The primary purpose of coordinated visualizations is to allow improved understanding through user interaction with different simultaneous representations of the data (Roberts, 2007). While choosing which combinations of views to use in order to generate insights can be complicated, several guidelines, including compactness and diversity of the visualizations, have been discussed in prior literature (Scherr, 2008/2009). It is hypothesized (Curry and Ross, 2015) that augmenting the traditional EEA approach with new analytic and interactive techniques will fundamentally enable new capabilities and insights to be derived from EEA, resulting in superior dynamic strategies for resilient systems. These extensions of the existing EEA framework enable the framing and analysis of large-scale problems, such as those posed by DoD’s Engineered Resilient Systems (ERS) efforts (Goerger et al., 2014). Recently, IEEA has been demonstrated in a case study on on-orbit servicing vehicles (Curry and Ross, 2016).

IEEA FRAMEWORK

IEEA leverages human-in-the-loop (HIL) interaction to manage challenges associated with the large amounts of data potentially generated in a study, as well as to improve sense-making of the results. By allowing the structured evaluation and visualization of many design alternatives across many different futures and potential lifecycle paths, this new approach enables the design of systems that can deliver sustained value under uncertainty.

DESCRIPTION OF IEEA FRAMEWORK MODULES

The purpose of IEEA is to "guide the...practitioner through the steps of determining how a system will deliver value, brainstorming solution concepts, identifying variances in contexts and needs (epochs) that may alter the perceived value delivered by the system concepts, evaluating key system trade-offs across varying epochs (eras) to be encountered by the system, and lastly developing strategies for how a designer might develop and transition a particular system concept through and in response to these varying epochs". To that end, as shown in

Figure 4, the IEEA framework is characterized by 10 individual processes that can be abstracted into six main modules:

1. **Elicitation** of relevant epoch and design variables (often through interview),
2. **Generation** of all epochs, eras and design tradespaces (often including enumeration),
3. **Sampling** of epochs and eras in which to evaluate design choices,
4. **Evaluation** of designs in sampled subset of epochs and eras
5. **Analyses** of design choices in the previously evaluated epochs and eras, and finally
6. **Decisions** of final designs based on iterative evidence from previous modules.

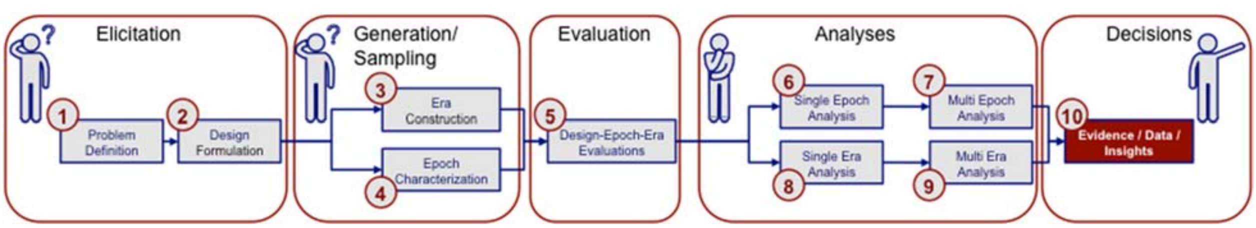


Figure 4 Framework Models and Processes

While the sequence of these modules flows logically, IEEA is intended to be an iterative process where users can go back and change responses within earlier modules at any point to reflect what they have learned from later ones. The six modules are composed of the 10 processes, but depending on the nature of the study and the type and fidelity of information available to the analyst, it is not strictly required that each process step be applied. Many of the techniques

discussed in Curry et al. (2015) can be applied to augment and facilitate a practical implementation of the workflow. For example, OLAP techniques may be applied to improve data handling, and search algorithms may improve our ability to offer more informed recommendations to decision-makers during the epoch-era analysis process. Similarly, enhanced human interaction techniques and visualizations may aid in the analyses of the vast amounts of information required to reach an informed decision.

CASE APPLICATION: COMMERCIAL OFFSHORE SHIP

This case applies the Interactive Epoch-Era Analysis (IEEA) framework on a case study from commercial offshore ship design, incorporating techniques from visual analytics such as interactive visualizations to gain insight from large, high-dimensional data sets resulting in improved strategies for value sustainment. New prototype visualizations are described which are motivated by a need to address design questions that are not well-suited for analysis with metrics, often applied in other EEA case studies, such as fuzzy Pareto number (FPN) or fuzzy normalized Pareto tract (fNPT) alone. For the offshore ship design case, this includes assessing the trade-off between designs optimized to target the primary mission versus being robust for uncertain subsequent missions. Further, considerations related to the implementation of interactive visualization applications, such as scalability and latency, are discussed emphasizing a need for continued research on methods for effectively handling large, high-dimensional data sets in design of complex systems under uncertainty.

This case study is based on a commercial ship case, developed by Rehn et al. (2016). A more detailed discussion of the case setup is provided in that paper. Offshore ships, in contrast to traditional deep-sea cargo ships, are designed to provide special operational services typically related to the offshore oil and gas industry. This group of ships comprises platform supply vessels (PSV), inspection maintenance and repair (IMR) and offshore construction vessels (OCV), to mention a few. A recent period of high oil prices and deep sea petroleum discoveries has spurred the development of offshore oil and gas fields. Thus, there has been a growing need for offshore services, including well maintenance and intervention services with light, riserless technologies. OCVs have taken an increasingly large part in the development of these, in particular for the marginal fields, due to their price competitiveness. Additionally, the Deepwater Horizon oil spill in 2010 in the Gulf of Mexico has changed some of the focus for the offshore shipowners towards being able to provide various deepwater emergency and rescue operations. This strong market period has characteristically driven the design of offshore ships towards multifunctional, gold-plated and expensive solutions (Garcia et al., 2016, Sep). However, the recent oil price collapse of 2014 has had a significant impact on the offshore markets, rendering many of these multifunctional ships less competitive against cheaper, specialized ships. The current situation in the offshore industry serves as a good example of the importance of focusing on value robustness and operational flexibility as key factors for success in a highly volatile maritime industry (Erikstad and Rehn, 2015; Garcia et al. 2016).

Offshore ships are usually built either for a specific long-term contract or on speculation. A long-term contract may last 5-10 years, and these ships are often specialized for the particular

mission. Ships built on speculation tend to be more multifunctional, to be able to take on different contracts. If these ships do not get any lucrative long-term contracts, they are often offered in the spot market to take on various short-term contracts. If a ship does not get a contract, it is idle for short periods or laid up over longer periods.

This case study motivates several questions, the evaluation of which may be aided using interactive applications described in this paper and by prior IEEA case studies:

1. What is the trade-off between optimizing for the primary contract and making the design robust to more than one contract in terms of the number of acceptable designs in the tradespace?
2. What is the impact in terms of both cost and reduced performance when attempting to ensure that designs satisfy all potential contracts?
3. What are the benefits and drawbacks of active versus passive value robustness?
4. Which contracts are most challenging to satisfy?

PROCESS 1: VALUE-DRIVING CONTEXT DEFINITION

The first process defines the stakeholders, problem statement, exogenous uncertainties and the basic value proposition for the system. The business opportunity for a new offshore ship design emerges from an expected strong demand for offshore oil and gas over the next couple of decades, despite recent short-term oil price volatility. The Deepwater Horizon accident has further resulted in an increased focus on being able to provide advanced offshore emergency services in the Gulf of Mexico. An offshore shipowner wants to target this business opportunity, and, in particular, a potential five-year contract for a large oil company. The shipowner wants to have a profitable and eco-friendly solution.

PROCESS 2: VALUE-DRIVEN DESIGN FORMULATION

The second process begins by defining the statements of needs, which become the attributes of system performance; along with utility functions describing the preference for each attribute. The system boundary for the single ship design is around the ship itself and does not consider, for example, the total profitability of the overall shipping company. Profitability is a measure of the ability of the design to generate profits, and eco-friendliness represents the ability of a design to reduce emissions during operation and transit. The non-monetary and monetary value attributes are kept separate due to their temporal differences in the model, which is further discussed in Rehn et al. [11]. In the model, profitability is considered at the era level, while eco-friendliness is considered at the epoch level.

Even though value focused thinking involves exploring various high-level solution forms, we assume the form of a standard single-hull OCV for demonstration purposes in this case study. The following ship-level design variables are considered: length, beam, depth, power, accommodation, main crane, light well intervention tower, moonpool, fuel type, dynamic positioning, remotely operated vehicle (ROV), pipe laying capability and design for changeability level.

PROCESS 3: EPOCH CHARACTERIZATION

In process 3, the key contextual uncertainties are identified so that epoch variables can be characterized. Based on the system boundary defined, eight epoch variables are identified, as illustrated in Figure 5. These epoch variables represent mainly the details of missions for a ship, operationalized through the contract rate and technical requirements, and the details of the operational area including the sea state and water depth.

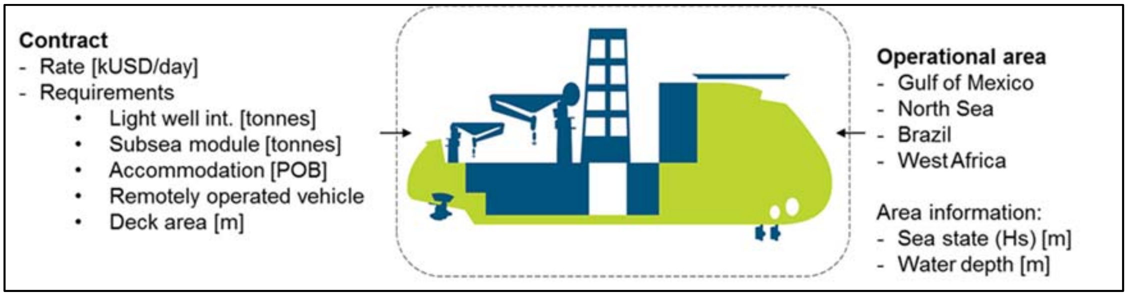


Figure 5 System boundaries and epoch variables (Rehn et al., 2016)

PROCESS 4: ERA CONSTRUCTION

This process constructs era timelines composed of multiple sequences of epochs each with a set duration to create long-run descriptions of possible future scenarios a system may encounter. Simulating lifecycle performance in this way allows an analyst to evaluate path-dependent effects that may only arise when uncertainty is time-ordered. The activities in this process are in many ways analogous to those used in narrative or computational scenario planning. The future timelines can be constructed manually with the aid of expert opinion (narrative) or by implementing probabilistic models (computational), such as Monte Carlo simulation or Markov chain models that define epoch transitions.

Three narrative scenarios are considered in this case study. In two of the eras the ship gets the targeted five-year contract initially, and experiences a relatively strong market the rest of the assumed 20-year lifetime. In the third era, the ship does not get the targeted contract due to a market collapse.

PROCESS 5: DESIGN-EPOCH-ERA EVALUATION

The first four processes defined the relevant elements of the models that will be evaluated in the fifth process. The previously defined models are integrated to map design and epoch variables into stakeholder value and expense. Many techniques are available for assessing both value and expense of a given system. A generalized approach may include multi-attribute utility (MAU) to quantify stakeholder value and multi-attribute expense (MAE) to quantify expense. This step connects the value space and the design space in a mapping correspondence, often via an intermediate performance space. For the offshore ship, the various key performance indicators are estimated based on simple relations from the design variables, including speed, deck area, dead weight, acquisition cost and operational costs. Designs that violate the technical requirements in an epoch are rendered infeasible.

PROCESS 6: SINGLE EPOCH ANALYSIS

Single epoch analysis is comparable to what is often referred to in practice as tradespace exploration. Within a given epoch, a scatter plot of cost (MAE) versus benefit (MAU) can be constructed that is fixed for short-run periods of stable context and needs (i.e., an epoch). Typically, decision-makers want to identify the frontier of Pareto optimal designs or, more generally, designs that are "close enough" to the Pareto front. Here the notion of "close enough" is operationalized through the Fuzzy Pareto Number (FPN) which is used to quantify the distance from the Pareto Front for each design in each epoch. FPN is a "within-epoch" metric and its value for a given design will change in different epochs. Decision-makers can gain insights regarding the difficulty of a particular set of context and needs by visualizing how points move in the design space as the epoch and FPN values change. Additional insights may be gained from interactively filtering the design, performance, or value variables. This can be performed with the aid of the filtering application shown in Figure 6 that allows the decision-makers to interact with their data to identify designs and epochs of interest. It also allows them to assign any of the defined variables to the radius, color or x-y location of the points in the scatter plot to explore the data in four dimensions and better comprehend the behavior of the designs.

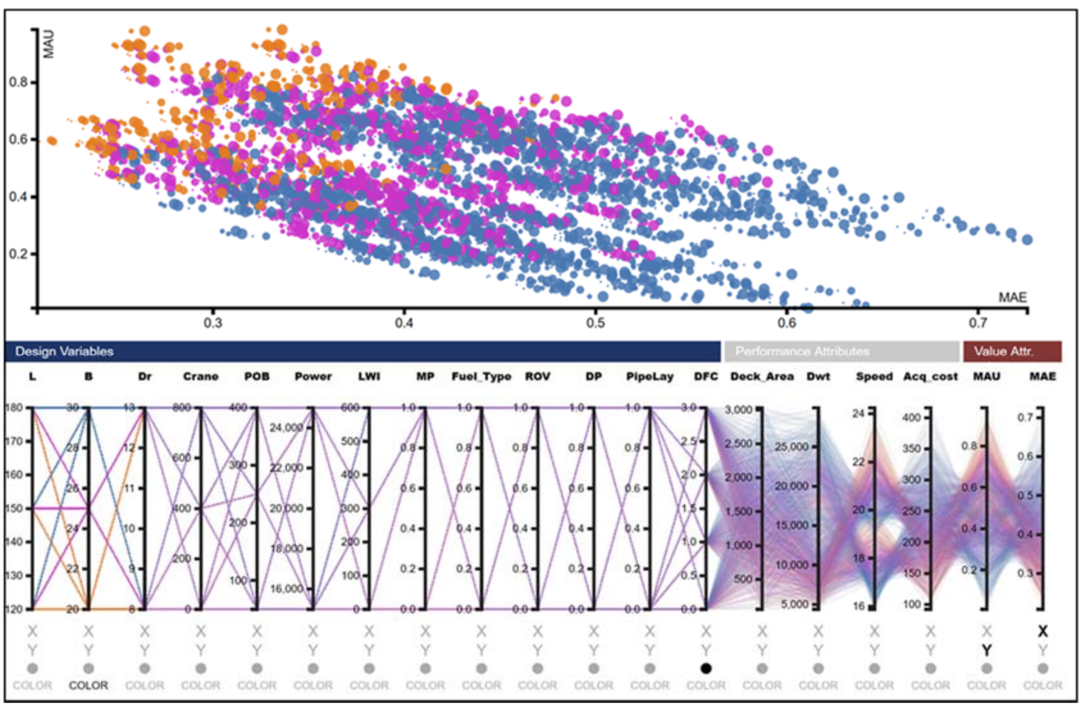


Figure 6 : Interactive Filtering Application for tradespace exploration for the offshore ship design base case

Figure 6 illustrates the tradespace for the offshore ship design base case that is the targeted contract with no technical requirements. Hence, at this initial stage, one can focus on understanding the dynamics of the underlying system. In this particular case study, the MAU only comprise one utility function, which is eco-friendliness, even though the figure indicates a multi-attribute utility function on a general basis. The interactive filtering can aid in visualizing

the exploration process of understanding the relative significance of individual design variables, as illustrated. For instance, filtering by beam and length, one can see that relatively slender ships tend to contribute to low FPN values. However, this again makes a design less stable in the water, which restricts the possibilities of retrofitting heavy equipment on deck without intervening with the main hull. Further, one can directly see the trade-offs of adding DFC levels, as design points shift right in the tradespace with increasing DFC due to increased cost.

PROCESS 7: MULTI-EPOCH ANALYSIS

The activities of process 7 allow decision-makers to gain deeper insights by evaluating metrics between and across epochs to gauge the impact of uncertainties on system value. This includes the evaluation of short run passive and active strategies for achieving value sustainment such that systems can maintain value delivery across different missions or changing contexts. A system that is passively robust is insensitive to changing conditions and continues to deliver acceptable value. Alternatively, a system that suffers deterioration in value due to evolving conditions may benefit from the use of change options that make it flexible, adaptable, or resilient.

<u>Evaluating Passive Strategies for Value Sustainment (Robustness)</u>

In general, we want to have the lowest cost ship that can fulfill the technical requirements of a contract. Since ships that do not have the required technical equipment for an epoch are considered infeasible, the number of designs in the tradespace as well as its shape will change depending on the epochs. Equipment is typically a large cost driver; hence, trade-offs are likely required between optimality in any one epoch versus how many of the enumerated epochs can be satisfied when using passive strategies only. The percentage of enumerated epochs satisfied at a given fuzziness level is quantified using the fuzzy normalized Pareto trace (fNPT) metric. A proper exploration of the trade-off between “closeness” to the Pareto front (FPN) and passive robustness across various epochs (fNPT) is important when extracting insights from these large, high-dimensional data datasets that are produced in the design process.

When examining this trade-off, attempting to look at all data dimensions of all possible designs across all possible epochs can be daunting for decision-makers. Even with clever visual encoding, visualizations that show all the data could likely incur additional cognitive load for the users rather than reduce it. Fortunately, depending on the task they are focused on, an internal mental representation of all data is not strictly necessary for an analyst. The interactive heatmap visualization shown in Figure 7 is one example of a simplified visualization that can show the compromise between Pareto efficiency (FPN) of designs within an epoch and the frequency with which they maintain that level of efficiency across multiple epochs (fNPT). This is illustrated in Figure 7 for the offshore case, and we can see that there are no designs that are Pareto optimal in all enumerated epochs. Accepting designs slightly away from the Pareto front or relaxing the constraint that epochs must be satisfied allows additional design candidates to be identified. Figure 7 shows that the fuzziness (e.g. threshold FPN value) needs to be relaxed to approximately 45% for any designs to be in the fuzzy Pareto set for an estimated maximum 87% of all epochs. This indicates that, in fact, no ship can satisfy all contracts and that the most

multifunctional passive ship can satisfy a maximum of 87% of the potential contracts requirements.

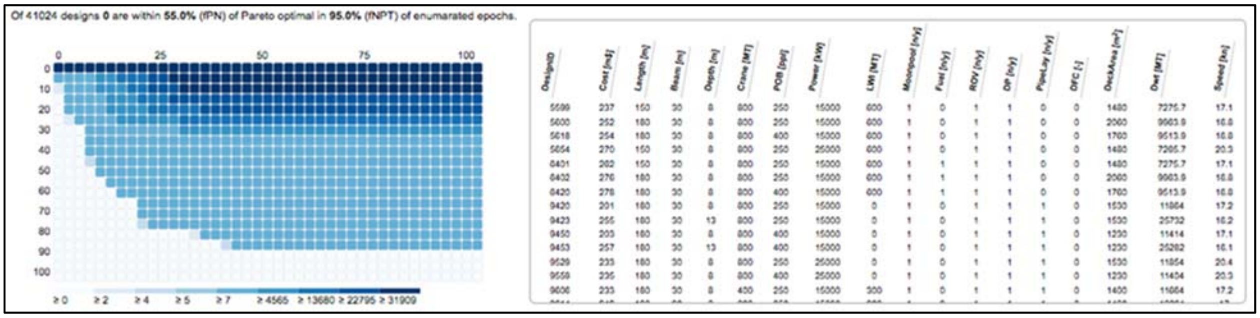


Figure 7 Interactive heatmap visualization (left), inspection table (right) for ships in selected tile in heatmap.

The interactive heatmap provides a high-level overview of trades between efficiency and robustness. But does not answer the questions: if an analyst wants to examine more complex trade-offs, for example, how restrictions on cost or other performance attributes impact the trade-off between FPN and fNPT, or, alternatively, if an analyst wants to identify whether certain epochs, stakeholders or context variables are more problematic than others for system value sustainment or they have a disproportionate effect on restricting the space of available alternatives. This type of information cannot be obtained from the heatmap visualization or aggregate measures like fNPT. More complex or nuanced questions like these require the examination of additional data dimensions that can be difficult to visualize and can also present added computational challenges.

This type of analysis is possible, however, with the aid of a more sophisticated visual interface like the example shown in Figure 8 , where a combination of online analytical processing (OLAP) and binned aggregation for fast filtering and interaction with larger data sets are applied. This visualization can also be easily scaled to case studies involving millions of designs and large numbers of data dimensions. This is possible because, rather than plotting every data point, each dimension is binned into a histogram that allows filters to be placed on individual data dimensions to see how that impacts the other dimensions in coordinated views. A list of candidate designs that match the filters is then displayed in the list on the right.

An analyst using this type of interactive visual interface can extract deeper insights about trade-offs by setting filters on various data dimensions to explore how those constraints impact other data dimensions or the list of available designs. For the commercial ship case study, this can be applied to gain a better understanding of the impact of fuzziness (FPN) and cost constraints. For instance, the designs that tend to be acceptable in most epochs, explored in the heatmap visualization in Figure 7, also tend to be among the most expensive in the tradespace. In fact, no matter how much the fuzziness threshold is relaxed, there are no designs that more than 85% of the epochs for a cost lower than $285 million. An analyst interested in achieving a lower target cost would need to examine in detail the cost savings that could be achieved by eliminating certain epochs (e.g. contracts, missions) which would result in a lower fNPT.

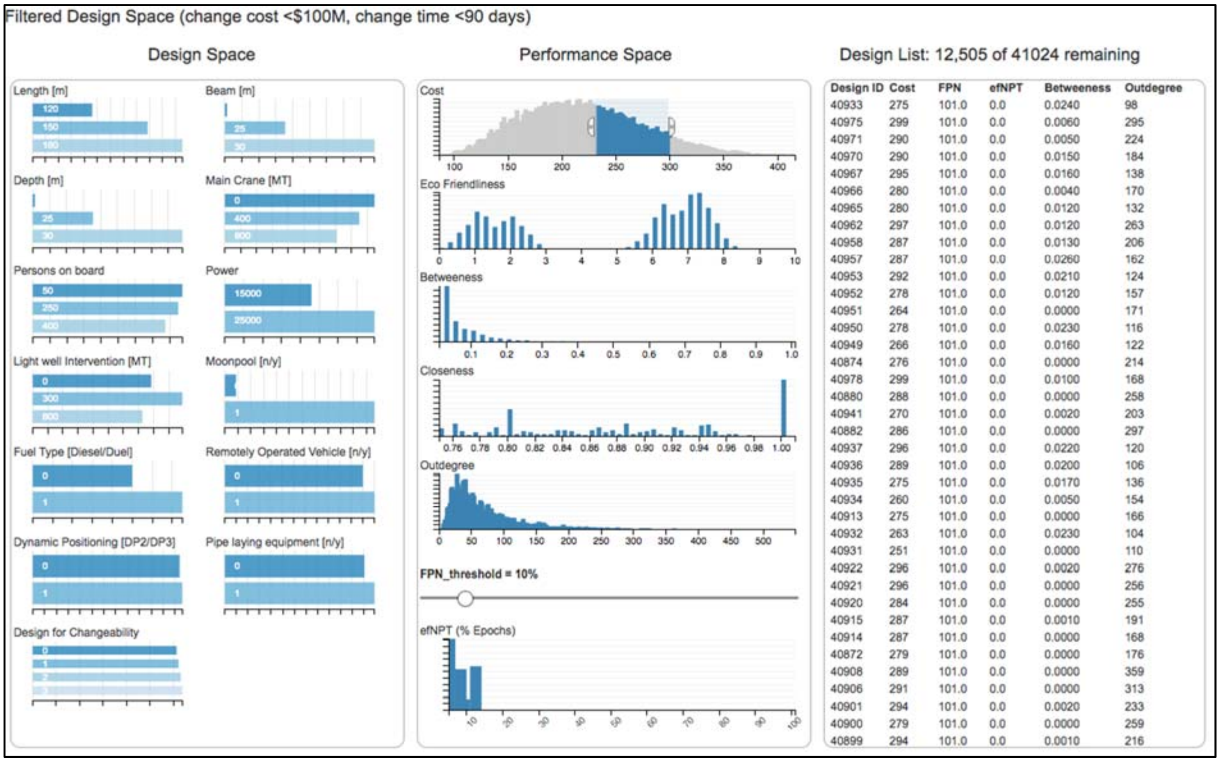


Figure 8 Interactive Filtering Application: implementing OLAP for the offshore case (Filtered for allowable change cost < \$100M and change time < 90 days)

PROCESS 8 AND 9: SINGLE AND MULTI-ERA ANALYSIS

Implementation of changeability in the offshore ship case enables the system to mitigate risk and take advantage of opportunities in a future operational context. This is enabled by initially optimizing for the targeted contract, but also providing the flexibility to be able to change the design later based on the next state of operation, which is uncertain at the initial design stage.

An offshore ship may be seen as a movable flexible platform that can carry equipment that enables the ship to take on contracts of various types. The size of the platform may also change through, for example, elongation, but at a higher cost and duration, compared to more traditional equipment retrofits on deck. Single and multi-era analyses using interactive visualizations as shown in Figure 9 can aid in the assessment of different classes of changeability for the offshore design case. For brevity, this analysis is not discussed in this paper, but the interested reader is referred to previous demonstrations in prior case studies using IEEA (Curry and Ross, 2015) for further details.

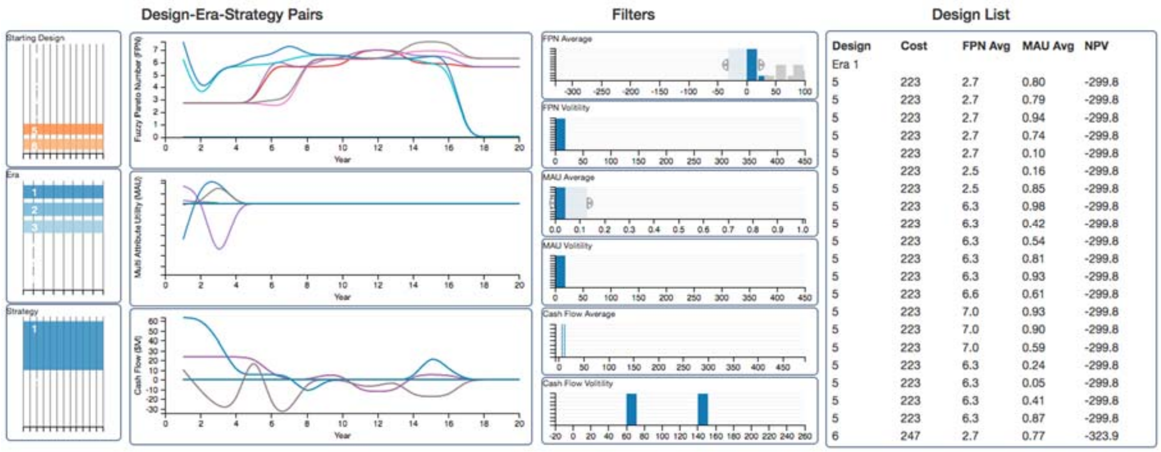


Figure 9 Interactive Multi-Era Application implementing OLAP for the offshore case

The analyses outlined in the preceding subsections provide a way for decision-makers to interactively evaluate the performance of multiple design alternatives across multiple futures. This creates opportunities for new insights at the expense of a potentially large and complex data set that can be difficult to analyze. The application of an interactive framework allows the users to visualize and engage with the data in new ways that can facilitate improved comprehension and decision-making. The insights that are extracted from this approach allow decision-makers to understand the characteristics of designs that can sustain value in all possible futures, through passive robustness or active changeability

IMPACT ASSESSMENT: HUMAN-IN-THE-LOOP SOFTWARE STUDY

A controlled human subjects experiment was designed to assess whether interactive visualization improves user performance for design problem tasks. A primary working hypotheses of this research is that the creation of visual analytic applications that couple interactive visualization with design methods (like Epoch-Era Analysis) will benefit human performance. To that end, this experiment required subjects to answer questions related to a simplified engineering design problem equivalent to a multi-epoch analysis problem. Subjects were randomly assigned to one of four treatment groups that were distinguished by the type of data representation or analysis tool they were given to solve the problem. It was anticipated that both the treatment group and individual differences between subjects impact performance as measured by task completion time and accuracy. Individual differences as discussed in this study refer to the subjects' personality traits and spatial reasoning ability. Results of the experiment confirm that this is indeed the case and that performance further relates to the task type in question. The question under consideration in this experiment is: *Does interactive visualization improve user performance for design problem tasks and, if so, what are the relative contributions of representation, interaction or other factor to user performance?*

There are two components to this question that must be tested. First, is there any measurable impact on human performance when individuals are engaged in performing analysis tasks commonly associated with engineering system design problems? Second, can the degree to

which interaction, visualization or other factors, such as individual differences in the users, affect human performance be isolated and identified. For the purpose of this study human performance is measured in terms of both the speed and accuracy with which a subject can complete a relevant task.

A primary working hypothesis of this research is that interactive visualization will improve design task performance for either, or perhaps both, of these metrics, and visual representation and interaction with data used to complete the task will impact performance in different ways. To decouple the relative contributions of representation and interaction a 2-by-2 factorial experiment with a total of 4 treatment groups was designed. Each treatment group corresponds to a different analytical tool that is provided to test subjects to complete a design task. All subjects were randomly assigned to a treatment group and asked to perform analyze a surrogate design problem, comprised of several tasks, that is a simplified version of a multi-epoch analysis problem. Because individual differences in personality or spatial reasoning ability may also play a role in task performance, data regarding these factors is also collected from participants using a pre-test.

A controlled human subjects experiment is an appropriate and effective approach for testing the working hypotheses. By controlling for whether or not a subject is given a data visualization and/or an interactive capability designed to aid them in solving a particular task the factors most influential to performance can be isolated. While the individual differences of the subject volunteers cannot be controlled this experiment can still effectively measure the impact of these factors by collecting a large and diverse sample set of participants. Further, participants are taken from two distinct cohorts to allow any factors unique to those groups to be measured and controlled for.

To provide a richer dataset and mitigate potential threats to external validity two cohorts of participants are evaluated in this study. The first cohort of participants was a selected pool of MIT graduate students. A second cohort of participants was drawn from volunteers using Amazon's Mechanical Turk (MTurk) online crowdsourcing marketplace. Data collected from the two cohorts is kept separate for data analysis purposes. The experiment is ongoing; in this current phase of IMCSE the second cohort has been evaluated.

Treatment Groups

The two primary controlled variables in this experiment were the presence or absence of (1) an abstracted representation of the data (e.g. chart, graph, visualization) and; (2) an interactive capability for manipulating or engaging with the data in some way (e.g. filtering, sorting). As discussed previously, it was hypothesized that these two things could aid a subjects understanding of a problem differently. To decouple the relative contributions of representation and interaction a 2-by-2 factorial experiment with a total of 4 treatment groups was designed. Each treatment group corresponds to a different analytical tool that is provided to test subjects to complete a design task. Figure 10 summarizes the four analytic tools subjects received as treatments in this experiment.

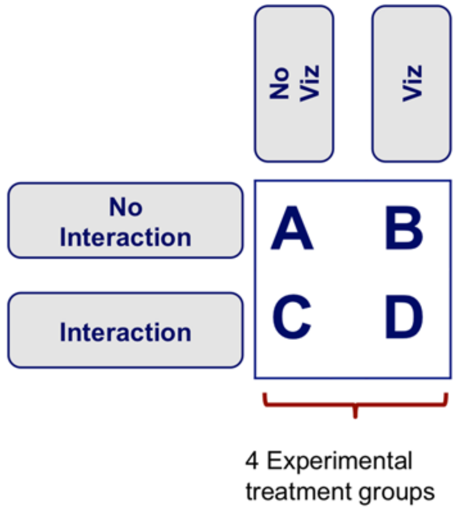


Figure 10 Experimental Treatment Groups

Subjects were randomly assigned to one of these four treatment groups, but each was given identical tasks and questions to answer.

Treatment A: Non-interactive Table

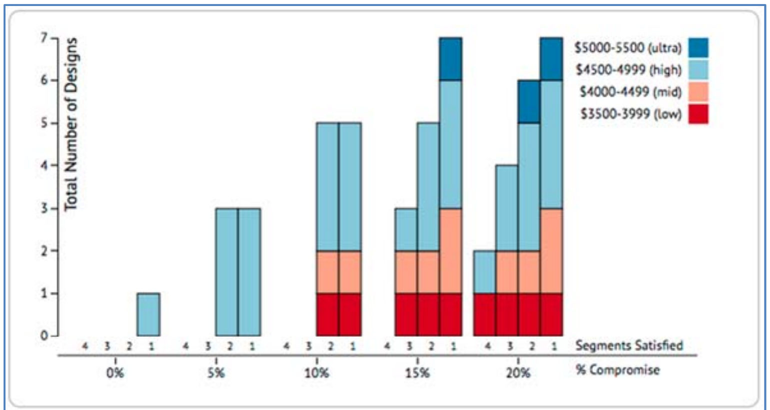
This group can be considered the control or standard treatment group in this study. Subjects are given a non-interactive table on the screen (Figure 11) that they must scroll through to determine the answer to task questions.

design ID	cost	reliability	mpg	top speed	0% compromise	5% compromise	10% compromise	15% compromise	20% compromise
1	4200	77.1	22.2	89	0	1	2	3	3
2	4050	76.2	20.4	92	0	1	2	3	4
3	4250	79	23.3	88	0	1	2	3	3
4	3500	60	27	81	4	4	4	4	4
5	3550	67.6	17.2	88	2	4	4	4	4
6	3700	76.2	18.2	82	1	2	4	4	4

Figure 11 Treatment A - Non-Interactive Table

Treatment B: Non-interactive Table + Visualization

This group extends the tool used for treatment A to include a non-interactive static graph. The graph is a stacked and grouped bar chart, as shown in Figure 12.



design ID	cost	reliability	mpg	top speed	0% compromise	5% compromise	10% compromise	15% compromise	20% compromise
1	4940	79	22.7	93	0	2	2	2	4
2	4260	75.2	18.9	88	0	0	0	1	1
3	4140	72.4	17.5	97	0	0	2	3	3
4	5270	77.1	29.4	84	0	0	0	1	2
5	3850	79	16.7	86	0	0	2	3	4
6	4660	90.5	22.7	87	0	2	2	3	3

Figure 12 Treatment B - Non interactive Table + Visualization

C: Interactive Table

This group extends the tool used for treatment A to include an interactive capability as shown in Figure 13. A row of input boxes above each column allows regular expressions (e.g. “<=20”) to be entered to filter the data set. Clicking on the column headers allows the data to be sorted by each data dimension.

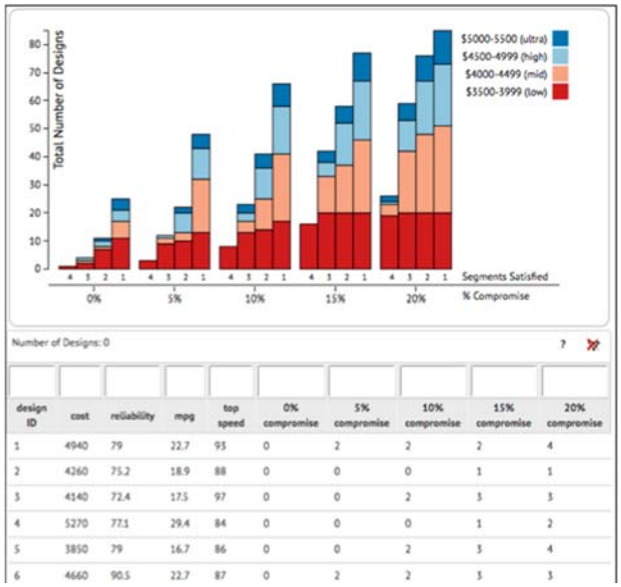
Number of Designs: 100 ?

design_id	cost	reliability	mpg	speed	0% compromise	5% compromise	10% compromise	15% compromise	20% compromise
1	4200	77.1	22.2	89	0	25	50	75	75
2	4050	76.2	20.4	92	0	25	50	75	100
3	4250	79	23.3	88	0	25	50	75	75
4	3500	60	27	81	100	100	100	100	100
5	3550	67.6	17.2	88	50	100	100	100	100
6	3700	76.2	18.2	82	25	50	100	100	100

Figure 13 Treatment C - Interactive Table

Treatment D: Interactive Table + Visualization

Treatment group D combines the interactive filtering and sorting capability of group C and an interactive bar chart similar to the static bar chart in treatment group B, as shown in Figure 14.



design ID	cost	reliability	mpg	top speed	0% compromise	5% compromise	10% compromise	15% compromise	20% compromise
1	4940	79	22.7	93	0	2	2	2	4
2	4260	75.2	18.9	88	0	0	0	1	1
3	4140	72.4	17.5	97	0	0	2	3	3
4	5270	77.1	29.4	84	0	0	0	1	2
5	3850	79	16.7	86	0	0	2	3	4
6	4660	90.5	22.7	87	0	2	2	3	3

Figure 14 Treatment D – Interactive Table + Visualization

TRIAL PROTOCOL

Participants were asked to complete a 3-part evaluation consisting of:

- The first part of this study asks participants to complete a standard test of spatial ability called the "Paper Folding Test"
- The second part of the study asks participants to complete a survey that evaluates the standard Big 5 Personality traits and locus of control commonly used in social science research studies
- The third part of the study asks participants to answer several questions about surrogate engineering design tasks using one of 4 possible web-application interfaces depending on which experimental treatment group they have been assigned into randomly. The possible interfaces are either:

  (A) a plain text table of data;
  (B) a static graph or visualization of the data;
  (C) a Microsoft Excel-like interactive table of data; or
  (D) an interactive visualization of the data.

Participants may exit the study at any point. Data from participants that exit the study early will not be included in the final analysis. The trial protocol is shown in Figure 15.

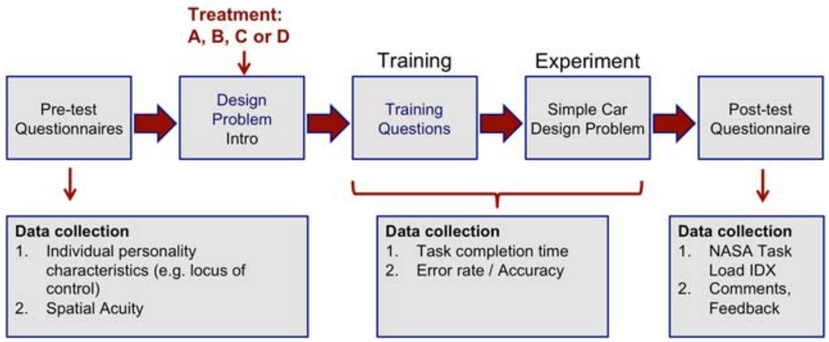


Figure 15 Trial Protocol

EXPERIMENTAL TASK

In the experimental task, participants were asked a series of 9 questions about the design problem they are given. The amount of data required to answer questions varied so that the impacts of overloading a subject's working memory could be assessed. Questions were defined by 3 different task types: (1) Filter tasks ask users to identify or count some subset of the data points; (2) Sorting tasks require data to be ordered numerically or alphabetically to answer the question; and (3) Trend task require the identification of a pattern in the data across groups or categories.

After completion of the 9 questions related to the surrogate design task the final question of section 3 asked subjects to rate their perceived workload while performing the tasks.

INTERIM RESULTS

This human-subjects experiment was conducted to decouple and evaluate the impacts of visualization and interaction on human performance when performing surrogate design tasks. The experiment is ongoing. The preliminary findings indicate:

1. Interaction seems to help with filter and sort tasks by reducing completion time and error rate.

2. Visualization seems to help with trend/pattern identification tasks by reducing completion time and error rate. Tasks that require analysis of larger amounts of data seem to make the benefit of either interaction or visualization seem more pronounced.

3. Individual personality differences also play a role in task performance. Spatial reasoning ability seems to correlate with task performance.

The expected research contribution is the experimental demonstration of the relative contributions of 2 primary components of visual analytic systems (representation and interaction) and how they are impacted by factors such as task type and individual personality differences. Further details of the study and findings will be available in (Curry, June 2017).

IEEA: Discussion and Future Research

The research presented applies the Interactive Epoch-Era Analysis (IEEA) framework, which provides a means for analyzing lifecycle uncertainty when designing systems for sustained value delivery. The framework has previously been tested on an on-orbit space vehicle case.

Application of IEEA to a case study for commercial offshore ship design in this phase of the research demonstrates key concepts and prototype interactive visualizations. IEEA extends existing frameworks with new analytic and interactive techniques that enable new capabilities and insights to be derived, which can lead to improved dynamic strategies for sustainment of system value delivery. In addition, these extensions enable the framing and analysis of large-scale design problems with uncertainty. Future work will extend this case study to include a deeper analysis of options at the epoch-level for changeability as well as era-level analysis of time-dependent aspects of system value.

The impact assessment experiment is ongoing, and results will be completed in the next phase of IMCSE research.

MODEL-CENTRIC DECISION MAKING STUDY

Models are increasingly used to drive major acquisition and design decisions, yet model developers, analysts, architects, program managers and senior decision-makers are faced with many challenges. Blackburn et al. (2015) captured many of these challenges in an investigation of the technical feasibility of radically transforming systems engineering through model-centric engineering. Digitized legacy systems and new digital system models will provide the basis for designing and evolving systems in the future (West and Pyster, 2015). This drives the criticality of models as assets and necessitates change in model-related policy and practices (Zimmerman, 2015). The Model-Centric Engineering Forum conducted by the US Department of Defense (DoD) Systems Engineering Research Center (SERC) in May 2016 fostered a dialogue between industry, government, and academia on current state of practice and vision for transformation (Clifford et al., 2016). Decision making will increasingly depend on models, yet there is relatively little empirical information on model-centric decision making.

As shown in Figure 16, there are three elements involved in model-centric decision making: the decision to be made, the digital thread/digital system model, and the human actors. The focus of this study is on the human actors.

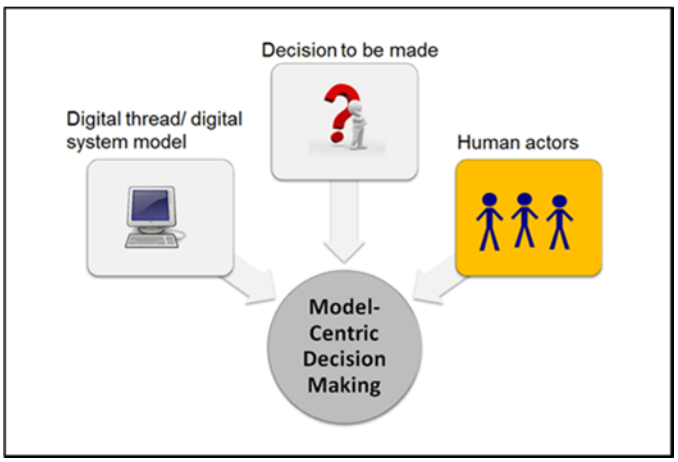


Figure 16 Elements of Model-Centric Decision Making

MOTIVATION

Ongoing research explores various dimensions of enabling model-informed decisions, as motivated by the increasing need for individuals and teams to make decisions based on models and model-generated information. Models represent an abstraction of reality in order to make predictions about the future, based on assumptions. Models can come in a variety of forms and formats, but fundamentally are an encapsulation of reality that humans use to augment their ability to make sense of the world, anticipate future outcomes, and make decisions. Among the many challenges are reasoning, comprehension and collaborative decision-making in the face of

uncertainty, combining artificial (model-generated) and real data, and effectively utilizing vast amounts of information.

The 2015 IMCSE Pathfinder Workshop validated the belief that improving human-model interaction would significantly improve model-centric engineering (Rhodes and Ross, 2015). Additionally, a 2016 workshop report sponsored by the National Science Foundation (NSF), the National Aeronautics and Space Administration (NASA), the Air Force Office of Scientific Research (AFOSR), and the National Modeling and Simulation Coalition (NMSC), highlights the need for understanding the individuals involved in the modeling process and how these individuals affect model development and usage (NSF, 2016). Central to this is the need to understand what engenders trust in models. While anecdotal stories of success and failure exist, empirical studies are needed to truly understand the many facets of human decision-making in model-centric engineering. For this reason, the research team initiated an exploratory study during this phase of the research, as highlighted in Figure 17.

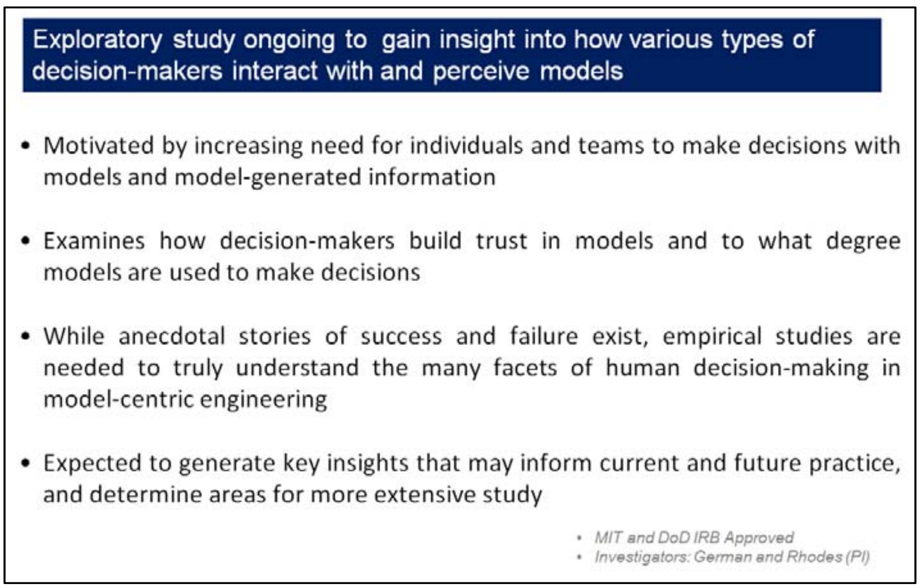


Figure 17 IMCSE Exploratory Study on Model-Centric Decision Making

INTERVIEW-BASED STUDY APPROACH

This study aims to generate insight into decision-maker trust and perception of models and model-generated information through expert interviews. Experts in system decision-making accumulate various kinds of knowledge and wisdom, often through years of hard-earned experience. Rather than theorize on how various actors interact with and trust models, this interview-based approach gathers qualitative, empirical data by asking them directly. This study is ongoing and is not meant to offer definitive truth for all types of decision-making with

models, but rather to serve as an exploratory study into a little-researched area. This study is primarily scoped to decision-makers and systems experts found within the defense community.

Unlike quantitative research, which advocates random sampling approaches, qualitative research seeks to “select ‘information-rich’ respondents who will provide you with the information you need” (Kumar, 2011). For this study, we have primarily used judgmental and expert sampling to identify “persons with demonstrated or known expertise in an area of interest,” (Kumar, 2011) along with individuals who, although perhaps not widely known as “experts,” were judged to have experience relevant for achieving the objectives of the study. In this study, we broadly view an expert as an individual who works or has worked as an actor within model-based decision-making processes, and can provide knowledgeable insight and perspective informed through his or her experiences. The definition of an expert is clearly open to interpretation as an “expert” may very well be in the eye of the beholder, and an improper interpretation on the part of researchers may lead to a biased sample of participants that fails to adequately represent a population. From our perspective, however, all participants were judged to have relevant experience and credentials through either their individually known work with models or that of the organizations for which they have worked, primarily experience found within the domains of defense and aerospace. While the study is ongoing, thirty individuals have been interviewed at the time of publication of this report.

Interviews for this study followed a semi-structured format that allowed interview participants latitude to share a wide range of perspectives and insights while following guiding questions aimed at generating insight for the study objectives. Table 1 presents a list of the general questions asked.

Table 1 List of Interview Questions

1. What types of decisions do you make, or help others make, with models?
2. What is the degree to which the decisions you make are based on models?
3. Do you view models as a primary or supplementary source in decision-making?
4. How do you develop trust in models?
5. How do you judge if a model can be trusted?
6. How much transparency do you desire?
7. What factors have led to inappropriate trust in models?
8. What limits your ability to use models to make decisions?
9. What challenges or failures have you experienced with the use of models in system decision-making?
10. What approaches or policies have been applied, or would you like to see applied, to mitigate those challenges?
11. How desirable would the ability be to directly interact with models real-time while making decisions?

DECISION-MAKING FLOW

High-level decisions incorporating an explicit model in the decision-making process include the following broad components:

1. A model that represents some aspect of the system of interest
2. Human actors
3. A decision to be made

While simplified, a generic conceptualization of the model-influenced decision-making process is helpful for facilitating discussion surrounding this research space. In this general framework, the information generated from a model is the common thread that connects the three generic commonalities listed above. First, a model must be created or already exist before it can be of use in a decision-making context – this creation of the model itself generates information relevant to decision-making before it is even "used." Next, the model in question generates information designed to facilitate a better understanding of an issue for which a decision must be made. Exactly what happens to this model information varies from context to context, but all contexts involve model information flowing *from* an actor (i.e. modelers or analysts directly interacting with the model), *through* another actor or actors, and ultimately *to* a final decision-making actor. Where decision-makers reside in the process seems to be more along a spectrum of the flow of model information. Within different decision-making contexts, actors may even find themselves in different roles. For example, in a mid-level decision-making context, an actor may be the individual *to* whom the information is flowing, yet in a higher-level decision, the same individual may become a *through* actor. In decisions involving more than one actor, however, all model-informed decisions involve information being generated and flowing from those directly interacting with a model, flowing through an actor or actors, and lastly reaching a final decision-maker to whom the information flows. To elucidate this conceptualization, it may be helpful to examine a specific case example that illustrates this flow of information.

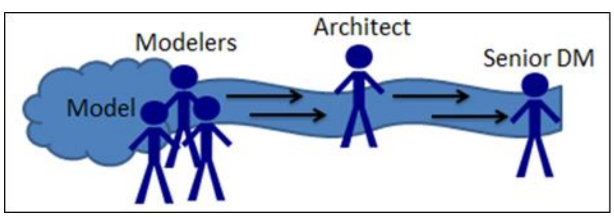


Figure 18 Figure 18 illustrates one such scenario where a model is used to inform decision-makers in a war game. The senior level decision-maker identifies a modeling need, and interfaces with a model architect to create the desired model. The architect works with a team of modelers who develop and test the model and produce model outputs that are communicated through the architect to the decision-maker in response to specific queries. In this case, the model information flows *from* the team of modelers who comprise the initial actors, then flows *through* the primary model architect, and finally *to* the Senior DM involved in the war game.

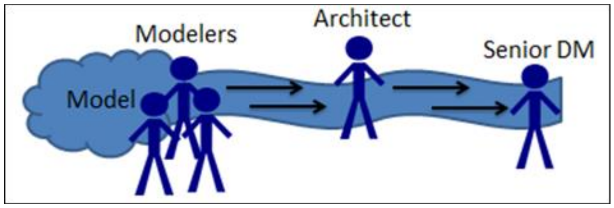


Figure 18 Three actor decision flow

At the end of the model-generated information flow there is a fairly discrete decision or set of decisions to be made. These high-level decisions, however, are influenced by countless smaller decisions and actions performed by various individuals within the flow of information. This study seeks to better understand the perspectives and thought processes of these various actors with the hope of better understanding the decision-making process as a whole. In the sampling process we sought perspectives from individuals from all three of our conceptualized categories; however, for the purposes of this paper, *through* individuals comprise the majority of participants.

TRUST IN MODELS

Ricci et al. (2014) describe how trust in models relates to a user's perception of how close to a specified reality a constructed model is perceived to be. Ultimately, a good decision "is one based on a trusted, truthful representation of both reality and values" (Ricci et al., 2014). The 2015 IMCSE Pathfinder Workshop report notes that numerous challenges exist within model-centric development, including challenges surrounding "perception of truthfulness and trust" in models, as this aspect of trust can ultimately affect "the timeliness, quality, and confidence in model-based decisions" (Rhodes and Ross, 2015). The Pathfinder report also expresses a desire not just for models to be trusted, but for that trust to be supported with underlying evidence. Blackburn et al. (2015) articulates a vision for developing model-centric environments into a "single source of technical truth" for decision-makers. West and Pyster (2015) communicate the idea of digital system models offering an "authoritative representation" of systems. Gass and Joel (1980) note, however, that all models "reflect modelers' views of how the decision problem can be resolved," and that these views carry inherent assumptions and limitations that decision-makers must consider prior to determining if the subsequent modeling results appropriately align with their decision at hand. With this in mind, the goals of developing single sources of "truth" and "authoritative data" will require decision-makers to evaluate and determine how much trust they should place in this data. This trust can be improperly calibrated, however, potentially resulting in overreliance or underutilization. Engendering an appropriate level of trust within decision-makers is crucial to effective use of models in decision-making.

Literature addressing human trust in automation offers insight that can be useful when applied to this discussion on human trust in models. This relationship seems rather natural when considering that automation may arguably be nothing more than a model of operation algorithmically programed into a machine. In the article "Humans and Automation: Use,

Misuse, Disuse, Abuse," Parasuraman and Riley (1997) highlight multiple potential pitfalls to consider when placing humans into interaction with automation. Misuse is defined "as overreliance on automation (e.g. using it when it should not be used, failing to monitor it effectively), disuse as underutilization of automation, […] and abuse as inappropriate application of automation by designers or managers". While examining factors that may contribute towards use and application of automation, Parasuraman and Riley identify that "trust often determines automation usage". This taxonomy of use, misuse, disuse, and abuse can provide a useful framework for thinking about how humans interact with complex models as well. But what exactly is meant by "trust?" Lee and See (2004) define trust as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability". Specifically addressing misuse and disuse, Lee and See express that "[o]vertrust is poor calibration in which trust exceeds system capabilities; with distrust, trust falls short of the automation's capabilities". This idea of calibration "refers to the correspondence between a person's trust in the automation and the automation's capabilities." Trust in automation implies belief that the automation will do what it is supposed to do, while trust in models assumes that the models will provide the information you want. Both automation and models represent technologies that require a certain amount of trust as the underlying processes and assumptions may be difficult to fully understand. The goal is not just for models to be used, but to be used appropriately; models, much like automation, have limitations of effectiveness and applicability. Overreliance in models can lead to misuse by inappropriately applying models outside of their inherent limitations. Conversely, improper lack of trust in models can lead to decision-makers discounting relevant model information that could have otherwise aided in the understanding and solution of issues. By examining the human aspect of human-model interaction, this study aims to generate understanding that can lead to appropriate "calibration" of human trust in models. Before seeking to influence the human actors, however, it is necessary to understand how those actors actually work in practice.

Consciously or not, decision-makers must have a certain amount of trust present before model-generated information is used in the decision-making process. Few of the actors interviewed have consistent processes to develop trust in new models, yet all have various factors they consider when determining trust. Some factors prove unique to specific individuals or groups of individuals along the flow of information, while other factors appear to be common for individuals throughout the entire flow. In addition to processes or factors influencing decision-maker trust, we want to know more about what specific attributes or types of information about models that decision-makers and actors care about knowing.

INTERIM FINDINGS

The model-centric decision-making study continues in the next phase (phase 5) of IMCSE research. Thirty interviews conducted in this phase (phase 4) have resulted in a set of interim findings on model-centric decision making. Eleven interim findings are described below. Further analysis and validation of findings will be performed in the next phase of the research.

TECHNOLOGICAL AND SOCIAL FACTORS INFLUENCING TRUST

As summarized by one participant, "trust is terribly important" within the modeling and decision-making process. While few of the interviewed experts have a specific process used in determining trust, every participant has various factors that they consider while determining the amount of trust to put in a model. This trust is also very contextually dependent, meaning that the trust is not so much in the model as an entity, but in the usefulness of the model for a specific decision at hand. Various factors influence individuals' trust in models, yet these factors may vary in importance depending on the specific individual involved. A clear theme that has emerged from the interviews, however, is that both technological and social factors come into play when determining the amount of trust that any type of actor is willing to place in a model. In many cases, the importance of technological factors appears to diminish in relation to social factors as actors move further along the flow of model information. Figure 19 illustrates various technological and social factors influence a decision-maker's trust in a model.

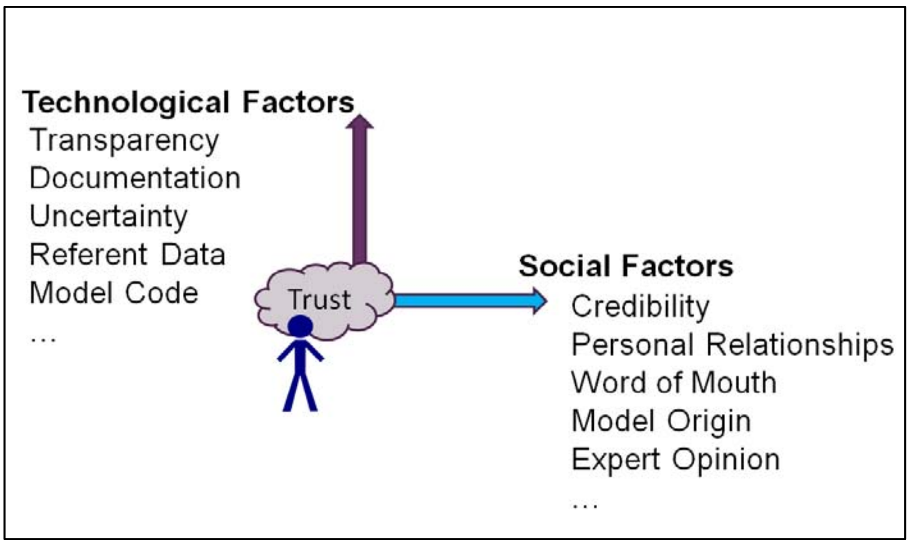


Figure 19 Technological and Social Factors Influencing Trust

The factors listed are not all-inclusive, but represent some of the factors identified through the interviews. While there may be trends in comparing important factors between the "*from, through, to"* categorizations of actors, such as the generalization that social factors seem more salient for *to* actors than for *from* actors, this is still dependent on the specific individuals involved. A strongly supported generalization, however, is that both technological and social factors play an important role in influencing an individual's trust, and any attempt to understand trust without considering both types of factors would be lacking.

IMPORTANCE OF COMMUNICATION

Communication, not surprisingly, arose as a key attribute of effective model decision-making. Before any effective modeling can be accomplished, senior level decision-makers must construct the problem statement clearly and in a form that unambiguously expresses the information they desire from models. Oftentimes the problem can change, however, therefore

consistent communication of the problem at hand is crucial for allowing individuals below them to create or use models to generate relevant and useful information. The onus for this specific communication does not fall solely on senior levels, however, and lower levels must actively update senior decision-makers on progress to gain feedback on whether they are addressing the actual problem. Senior decision-makers must likewise be open and available, to the extent possible, to provide this feedback as necessary. As noted by an interview subject, "models […] bring their own language with them" that can create communication barriers that stifle decision-makers' understanding of the model output presented to them. Unless a decision-maker is similarly an expert in the model, there needs to be a "translation between output to decision-maker speech," before the information can usefully be incorporated by the final decision-maker. Modeling aims to provide an asset for a decision; however, this asset cannot be effective if it is not useful for a decision-maker, and it cannot be useful if not understood. Instead of relegating discussion between actors to the beginning and end of a decision-making process, employing continuous and iterative communication may further reduce the acceptance barrier by allowing decision-makers to feel as if they walked the up-stream actors to the final model outputs. The flow of information between actors, including both expression and interpretation of information, must be intentional and unambiguous.

Transparency

Most of the interviewed modelers, analysts, and architects emphasized the importance of having access to precise technical information of models, oftentimes stating a desire to have access to code and the "guts" of the model in question. One such practitioner expressed that he "hopes everyone wants full transparency," seeming to assume that the desire for full transparency is a given for anyone making decisions from models. Transparency serves to enable an understanding of how a model actually works in order to determine if the model should be used for a specific decision. The understanding of a model encompasses, but is not limited to, a model's code, and transparency should include access into practices and decisions involved in creating and validating the model. Moving further along the flow of information decision-making, however, precise information about the models may become less desired, and even unwanted. Comments such as "I trust the people below me" convey the paradigmatic shift that occurs. While details such as model assumptions and uncertainties remain desired, the need for intimate technical knowledge seems to fade. Responses suggest that, even if an actor does not personally require full transparency into a model, transparency should still be available to trusted actors before them in the flow. This suggests a significant point: as actors move further along the flow of information and have less time and ability to personally investigate a model and build their own trust in the model, their trust instead shifts more onto their people to investigate the model for them. In this understanding, the trust for decision-makers is "implicitly on the models, but explicitly on the people."

Understanding of Assumptions and Uncertainty

All models are inherently abstractions of reality that contain assumptions and uncertainties. Models are created for a specific reason and context, and while the assumptions within the model aim to help answer those questions, they also fundamentally create bounds of model

applicability. Failure to properly understand the inherent limitations found within a model increases the likelihood of the model being used inappropriately. As models cannot perfectly encapsulate and relate the situation of interest, uncertainty is fundamentally a part of the results, and uncertainty is also fundamentally a part of determining if the results are appropriately relevant to the decision to be made. This uncertainty must be sought after, understood by the sources of model information, and then passed clearly along the flow of information. There is a fundamental need to understand and express model uncertainties throughout the decision-making flow. Organizational and social dynamics can hinder this expression of uncertainty, however. In some instances, uncertainty about an answer may entail negative stigmas and imply failure to do one's job correctly. Decision-making cultures need to strive to drive out fear of uncertainty expression and transparency.

EFFECTIVE MODEL DOCUMENTATION

Model developers internally carry within themselves the most intimate knowledge of a model's limitations and capabilities. Similar to how modeling is a process of making the internal mental models and expertise found within individuals explicit, documentation is a process of making the assumptions and limitations of a model explicit. Models may very well be validated, even accredited; however, this validation and accreditation are for specific conditions, outside of which the model is no longer valid. Multiple interviews revealed the danger of assuming a model can extend to any context needed when in fact its appropriate contexts of use are much more limited. For a model to have any sort of reuse capability, these assumptions and limitations should be documented in an accessible way so that others can understand how they might appropriately apply the model to their specific situation. Models are built to answer a specific question or set of questions, and the early conceptualizations (e.g. whiteboard drawings) of the model and decisions made in the development process can provide important insight into understanding the model in addition to the documentation of assumptions within the model itself. These conceptualizations, if captured, can provide useful artifacts in the understanding and trust of a model. As models become more complex, documentation of assumptions and capturing of conceptual artifacts and decisions will likely prove crucial in allowing actors to appropriately calibrate their own understandings and mental models of if, and how, a model should be applied to specific decision-making scenarios.

PRIMARY VERSUS SUPPLEMENTARY

Of the experts interviewed, distinctions emerge concerning the primacy of explicit models in the decision-making process. Some view models as clear primary sources in decision-making, others adamantly express that they should only be supplemental sources, and still others present the oft-favored viewpoint of systems engineers – it depends. Those that favor models as a primary source in decision-making point to the benefits of increased knowledge and insight that models can provide if done correctly. Others that advocate for supplementary use emphasize the danger of abdicating the decision-making process to models, and point to the inability of models to capture every relevant factor in a decision. One participant noted an increasing reliance on modeling and simulation (M&S) in decision-making, unfortunately accompanied with the increasing desire to rely on M&S without having to "understand the fundamental processes behind it." The variations in responses serve to validate the non-

definitive (yet still insightful) answer of “it depends.” Truly, how models are viewed and used is dependent upon the model users and decision-makers, along with the modeling and decision-making context. Well-established and validated physics-based models, for instance, might prove to be a primary source in a decision-making scenario, while descriptive or predictive models that are less conducive to traditional validation may contribute more of a supplemental input within a wide range of other inputs.

INDEPENDENT REVIEW

Although models strive to reduce complexity of reality to understandable and workable abstractions, they can still be very complex. Verification and validation (V&V) are crucial for determining the efficacy and relevancy of a model for decisions (Pace, 2004). Just as skill is needed in model development and use, checks like V&V are required to hopefully catch the inevitable errors. However, effective V&V likewise requires skill and is liable to its own errors. One longtime system architect we interviewed emphasized the importance of utilizing independent experts who can review and render judgments concerning the credibility of results and believability of the data used. Such a team would be composed of individuals with areas of expertise relevant to the problem. One might view the team as analogous to a forensics team that closely examines the data and code being used and makes judgements that assess the efficacy of decisions made along the flow of model information. Depending on the model and decision-making context, the format and formality of reviews could range from formal, externally-based reviews, to informal, internal peer reviews within a team. Whatever the format, a form of review can serve an important part in the creation of an effective model, and as such, should be a process that is transparent to the decision-makers who are ultimately affected.

INVESTMENT BIAS

One individual related the story of a program that involved significant investment in modeling and simulation. When the time came for program decision-makers to make a decision, “they had no choice but to accept” the model’s answers “given the resources that were spent.” Such a story brings to light the potential bias that investment of time and resources into model development will yield correct and reliable results. Further interviews also revealed a potential for decision-makers to use money as a basis for establishing trust in model results. Money may sometimes offer a useful indicator of model capability; however, no matter how much money is spent on a model, the model is still bounded by the problem space it was designed to solve. Just because large amounts of money were spent on a model does not mean that it is appropriate for the decision at hand. If this issue is not a bias in some cases, then perhaps it may be a political pressure to make a decision based off the model results because of the money spent on model development – if not, the money was wasted. Such a logical fallacy should be countered by a fundamental term of economics: sunk cost. Once money and resources have been spent (sunk) they are gone, and no longer should have any bearing on decisions seeking to promote benefit in the future.

CONFIRMATION BIAS

In the words of one respondent: “Quite often, what I see is that decision-makers use models as confirmation bias.” This statement reflects one potential pathway for models to be used

inappropriately, namely, as a means to further one's own preconceptions or agendas that may be incorrect. Just because a decision-maker's intuition for a solution matches up with a subsequent modeling result does not mean the intuition or the modeling was wrong; in fact, it could be a testament to the decision-maker's experience. However, a senior modeler noted the challenge of guarding against bending a model and results to produce answers desired by decision-makers. Another participant expressed the "amusing thing" that in high-level war game simulations, the war games "almost always" are eventually modified so that your side wins. These interviews reflect the importance for all actors to honestly seek truth while participating in the modeling process. Modeling aims to provide solutions to problems; however, if generated and used to advance one's agenda or to inappropriately confirm preconceived notions, the "solutions" provided may in fact be more damaging than if models were not used in the first place.

ENDOGENEITY OF THE HUMAN

Underpinning this study has been the clear and consistent theme expressing the endogeneity of the human in the model-centric decision-making process. Many senior decision-makers do not have the bandwidth, training, or time to become technical experts in the models that are used to inform their decisions. How do they trust complicated models? As one senior-level decision-maker put it: "The answer is they trust the people." They trust that the people before them in the model information flow handled the data correctly, created, tested, used and analyzed the model correctly, and expressed the results accurately with appropriate information on uncertainties, assumptions, and limitations. Decision-makers trust that those individuals have the appropriate expertise and capability to understand and address the problem at hand. On the other hand, senior decision-makers also need to have the technical judgment to be able to "sniff out" the wrong answers, and have a healthy technical competence appropriate to the decisions being made. As systems and their models become more and more complex, the need for skilled and experienced individuals to work within the flow of information seems to be more necessary than ever. Yet the inevitability of aging and retirement guarantees that the experts of today cannot be the experts of tomorrow. Without the right people capable of handling the complexities we are creating, the system will fail, regardless of the technology and innovation we throw at it.

REAL-TIME INTERACTION WITH MODELS

A final question asked in the interviews concerned the desirability of being able to directly interact with models in real-time while making decisions. Overwhelming, the respondents view interactivity with models as highly desirable. After all, many decisions involve asking "what-if" questions about the model, and direct interaction could serve to gain insight, build intuition, and speed innovation without needing to go through other human actors. This support for model interactivity also comes tempered with caution from some individuals, however. Specifically, caution against allowing actors interactive access to models without a calibrated understanding of the model's capabilities and limitations. As related by one individual, in situations without this appropriate understanding, "I can get lots of results real quick, and I can make lots of bad decisions real fast." These interviews make abundantly clear the importance for properly understanding a model and its associated assumptions before determining one's

trust and usage of model results. Such an understanding is crucial for effective and appropriate interaction with models. So while direct interaction with models may be rightly desired based off its potential benefits, development and deployment of interactive models must also advance in a smart and conscientious manner to ensure that actors are not being set up for failure due to ignorance of their own limitations.

PRELIMINARY HEURISTICS (GUIDING PRINCIPLES)

A desired outcome of model-centric decision making study is a set of guiding principles, or heuristics that will be useful to practitioners and designers of model-centric environments. In this phase some proposed heuristics have been developed, stemming from literature review and the interviews. These are being further developed and validated in the next phase of research (German, June 2017).

The following are preliminary proposed heuristics:

USING MODELS IN DECISION MAKING

- Models do not have agency, and therefore do not have responsibility. The responsibility of decisions must be upon the human.
- Human are central to success of modeling and decision-making; models should be designed for humans, rather than the human being forced to adapt to models.
- A fundamental goal of modeling is to ensure the model is used appropriately given purpose and context.
- Decision-makers are much more likely to use a model's results if they feel like they walked the modelers to the solution.
- Model-centric engineering includes making sure the human actors are appropriately qualified to handle the model's information.
- Model design has inherent effects upon user behavior – model developers should have a sense of accountability that their decisions will affect user behavior, and should not expect users and decision-makers to adapt to the developer's view of "rationality".
- Beware the "ironies of modeling" -- improper models or improper decision-makers using models, can make the decision-making process worse.

IMPORTANCE OF MODEL CONTEXT AND ASSUMPTIONS

- Models are created for specific reasons and contexts, those assumptions fundamentally bound the model's applicability.
- *All models are wrong, but some are useful*, and before any can be useful, their limitations must be understood.
- Never assume a model applicable in one context will be applicable in another context.
- Independent review of models provides value through raising important questions, challenging assumptions and detecting biases that modelers may fail to see themselves.

- Documentation is a process of making the assumptions and limitations of a model explicit – without documentation, a model is not reusable.
- Not all models and modeling contexts are created equal – some models can be viewed as primary sources in decision-making, however, others should be viewed as more supplementary.

MODELING PROCESS AND DECISIONS

- Modeling is fundamentally a sociotechnical process composed of human actors interacting with models, model-generated information, and one another.
- The value of the modeling process stems from the flow of information through and between human actors involved in making and supporting decisions.
- Model-centric engineering may imply a transformative shift, but does not preclude learning from past experiences with the process of modeling.
- Money and time spent does not necessarily translate into a model that meets your situation.

TRANSPARENCY AND TRUST

- Engendering an appropriate level of trust within decision-makers is crucial to effective use of models in decision-making.
- Trust is a sociotechnical construct; you must examine both the technical and social factors if you want to understand trust in the model-centric context.
- Transparency is important to everyone, however, it also means something different to everyone.
- While the opportunity for full transparency is always desired, full transparency is not necessarily always desired – especially for higher level decision-makers.
- For those without an intimate understanding of the model, trust is implicitly on the models, but explicitly on the people.
- Model trust is a process of determining model applicability and efficacy for a decision at hand.

MODEL-CENTRIC ENVIRONMENTS

- Model-centric environments need to apply user-centered design -- designing for how the user actually performs, and not just what the user wants.
- Modeling environments need to be structured to minimize potential biases and failure modes that users are often unaware of themselves.
- Increased efficiency may also decrease time spent analyzing a problem, which in turn increases chance of poor judgement and biases.

Model-Centric Decision Making Study: Discussion and Future Research

The increase we see in system modeling is driven both by a desire to better understand complex systems and issues as well as by increases in technological and computational capability. Similar to technical modeling in many ways, automation involves increasing automation in systems as advancements in technology allows. Often this increase results in gains of efficiency and safety, yet the history of automation has also shown that humans are not just outside users of systems, but rather are endogenously critical components of the system. Experience has also shown that increasing technological capability for the sake of technical achievement, without proper consideration for the human component, can have dire consequences. Bainbridge (1983) writes about the "ironies of automation," where introducing automation can sometimes increase the workload and complexity of tasks it aimed to reduce. With gains in modeling complexity and capability pointing to a model-centric paradigm of engineering, we should be cognizant of potential "ironies of modeling" where failure to appropriately account for human decision-makers and actors results in worsening decision-making processes we aimed to improve.

This study aims to generate empirical insight into how human actors interact with and trust models, while also providing a starting point for continued exploration into how human actors and decision-makers trust, perceive, and interact with models. Through the interviews conducted, important considerations are identified surrounding human-model interaction and trust that experts deem important for effective model use and decision-making. These considerations include practices that interviewed experts implement to aid in their decision-making, and the identified challenges that can degrade effective model-centric decision-making (along with potential mitigations to challenges). The insights gained from these interviews are planned to be coupled with empirical case studies examining human interaction with complex, abstracted systems to gather information about how human actors and decision-makers actually perform in practice. The descriptive insights gained through empirical research will be bolstered with normative research on decision-making and biases. Taken together, we envision these various threads of research weaving together towards prescriptive outcomes of heuristics and design principles to inform policy, design, implementation, and use of model-centric engineering practices and environments.

RECOMMENDATIONS FOR FRAMING MULTI-STAKEHOLDER TRADESPACE EXPLORATION

Tradespace exploration is a rapidly advancing design and decision support paradigm that is particularly applicable to complex systems with many value-driving dimensions. These systems commonly have multiple stakeholders that can exert critical influence on the system's conceptual design, necessitating the satisfaction of their needs, and often requiring negotiation. Previous research has suggested that classic tradespace exploration activities may reinforce negative negotiation behaviors through their framing of the multi-stakeholder problem. This section presents active research in recommendations for supporting inter-stakeholder and stakeholder-data interaction, both fundamental to human-model interaction. These recommendations include the reframing of standard tradespace activities and visualizations using the combined insights of the negotiation, framing, and TSE literature and extend from problem formulation through exploration of the data.

INTRODUCTION

As modern engineering systems increase in size and scope, it has become increasingly necessary to consider the perspectives of multiple stakeholders in the conceptual design process (Garber et al., 2015). Stakeholders most commonly enter the design process as the definers of value – the desired attributes of the system and the reasons for which it is being designed. Many methods for approaching the multi-stakeholder problems choose to aggregate stakeholder preferences, reducing the dimensionality of the problem and providing powerful leverage for algorithmic design and optimization. However, these methods are only mathematically rigorous under specific axiomatic conditions and are by definition a simplification: often underrepresenting the true complexity of the problem (Scott and Antonsson, 2000). Additionally, and perhaps even more importantly, many stakeholders are reluctant to abdicate their decision-making authority to a model and therefore may reject normative frameworks for combining value functions. When stakeholders can exert influence on the design process up to and including potential veto power, the multidimensional comparison of *each individual stakeholder* is necessary in order to identify alternatives that are both "good" at their intended purpose and "fair" in accordance with the social dynamics between the stakeholders. Designs that lack either of these qualities may find themselves useless or unable to generate the buy-in necessary to continue with, and complete, detailed design and eventual operations.

This section will discuss the ability of tradespace exploration (TSE) and, specifically, multi-stakeholder tradespace exploration (MSTSE) to support early conceptual design of engineering systems with multiple stakeholders. MSTSE has been developed to target design tasks with stakeholders who are unwilling or unable to fit their preferences into a shared normative decision framework but who remain involved in the design process. Framing has been identified as a challenge leading to counterproductive negotiation tactics by previous MSTSE research, but a challenge that is capable of being ameliorated through creative redirection of attention and emphasis on group-dynamic data over individualistic data (Fitzgerald and Ross, 2015). Those

results are supplemented here with recommendations for framing adjustments throughout the MSTSE process, including early in the problem formulation.

MULTI-STAKEHOLDER TRADESPACE EXPLORATION

Tradespace exploration is a design paradigm that uses the analysis of many alternatives in order to build understanding of the tradeoffs between value-driving attributes that are available to the designers (Ross and Hastings, 2005; Ross et al., 2010a). Without restricting attention to a particular implementation, generally a TSE project will follow a procedure similar to this:

1. **Problem Formulation** – the structuring of the problem and scope of decision making. This normally includes the definition of the design space used to enumerate potential system alternatives, the context in which those systems will operate, and the stakeholders and value attributes used to assess them.
2. **Modeling/Evaluation** – the development and use of models for the purposes of evaluating the designs. Models can take many forms, which necessitates a selection of modeling technique(s) appropriate to the problem formulation. Creating models is itself nontrivially difficult and normally takes considerable effort without the benefit of reuse of previous models.
3. **Exploration/Analysis** – the attempt to curate insights from the model outputs. Stakeholders and analysts are both capable of performing this step, with different strengths and weaknesses. Exploration is typically intended to generate results capable of justifying a decision to select a given design alternative.

This knowledge-building process is particularly useful when applied to complex systems for which designers or analysts may not have a strong intuition of the dynamics at play. The presence of multiple cooperating or competing stakeholders is one such complexity. Early attempts to incorporate multi-stakeholder analysis into TSE simply used a value model for each stakeholder and used analysts to find design alternatives that satisfied each stakeholder's model. We refer to this type of analyst-driven exploration as "informal" MSTSE, to indicate the unlikeliness of reaching a formal agreement using the tradespace without stakeholder participation (not to imply any sloppiness in the construction or exploration of the tradespace). Informal MSTSE has the advantage of being able to be conducted by experts in a manner similar to most systems engineering activities, with the resulting lessons and insights then communicated to stakeholders before they engage in the "formal" negotiation or decision making process. However, this approach naturally risks costly iteration, as the negotiation may raise new questions that must be sent back to the engineers responsible for tradespace analysis and delay the final decision.

This weakness inspired a new approach consisting of parallel exploration of the data by each stakeholder, with the goal of uncovering emergent insights in the intersection of their exploration and facilitating dialogue amongst the stakeholders that could result in iterative refinement of their value model *during* exploration rather than separate from it (Ross et al., 2010b). Though effective at its intended purpose, this type of multi-stakeholder analysis was conducted entirely with the mindset, supporting visualizations, and metrics of classic TSE.

Efforts to formalize the concept of multiple stakeholders engaging with tradespace data into MSTSE sought to re-examine the latent assumptions in these methods, in order to confirm or reject their suitability for the additional complexity inherent in the multi-stakeholder problem (Fitzgerald and Ross, 2014). Framing was identified as a potential key roadblock to effective MSTSE, due to the aggressively individualistic framing of traditional, single-stakeholder TSE analysis leading to misplaced reference points for decision making and misattribution of gains and losses.

MACRO FRAMING AND MICRO FRAMING

The concept of framing has been used in many different ways, to describe many different ideas. In their most basic sense, all the uses of framing share one key feature: the understanding that contextual factors impact human perception and thus human action. The wide scope of framing can sometimes lead to confusion when discussing its implications. An instructive division of the relevant literature is by whether the framing occurs outside or inside the boundary of a specific case, which we call *macro* or *micro* framing issues, respectively. To illustrate the differences between these two types of framing, the following subsections will cover some of the prominent literature in the topics, following that with early research returns on the impact of framing in MSTSE.

Macro framing. Macro framing lies outside the domain of any single decision problem and deals with issues of writ-large beliefs and perspectives. Perhaps the most famous science-oriented discussion of framing is that of Kuhn (1962) on the subject of scientific revolutions. Kuhn describes the progress of science as one of prevailing paradigms that are upset by revolutions in favor of new paradigms. Revolutions are often characterized by heated debate between the scholars of the different paradigms, who frequently have difficulty communicating because they are figuratively speaking different languages. The paradigms can be viewed as frames (or perhaps lenses in this analogy) through which people ‘see’ an issue. Competing paradigms can make normative arguments in completely different directions, as the norms to which they appeal do not necessarily align. This can affect negotiations even at a mechanical level. For example, there is evidence that *differences* in outcome goal orientation and process goal orientation, two types of mental framing positively correlated with high-value negotiation results, can negatively impact the quality of negotiation outcomes. When measuring each type of goal orientation present in negotiations, similar levels of both key types of goal orientation resulted in better negotiation outcomes than simply having high absolute levels of goal orientation (Katz-Navon and Goldschmidt, 2009).

Schon and Rein (1994) also approach the issue of interpersonal conflict through framing, specifically targeting the realm of policy creation. They stress the importance of “frame reflection”: deliberately considering the differing frames of each actor as a preliminary step to effective policy design. Moreover, each actor can balance multiple frames, both rhetorical and action-oriented, that operate on different levels. At the highest level, “metacultural frames” are heuristic frames with highly engrained societal norms. They provide an example of the use of metaphors like “sickness versus health” to justify actions such as urban renewal, which may be

in direct conflict with a frame of “family and culture” on the same issue. These metacultural frames influence lower level “institutional” frames (general norms and actions of an institution) and down to “policy” frames (the framing of a particular issue). Though couched in the language of policy, due largely to the prominent role that ideology plays in political debate and policy creation, Schon and Rein’s work can be applied to any field with multi-party conflict over issues more fundamental than objective fact. Frame reflection allows participants in the conflict to examine not only where their own beliefs come from but also those of their counterparts. Though Schon and Rein rightly acknowledge the risk of relativist paralysis (e.g. questioning the objective validity of norms can lead to failure to act), they provide many examples, though not directly engineering-related, of frame reflection by key actors resolving entrenched conflicts by clarifying the decision criteria for each party to the other.

More generally, “macro” framing is often a subset of personal philosophy. Of particular interest to negotiation is the issue of fairness or equality, as it has considerable bearing on the evaluation of outcomes in group problem solving. Raiffa (2002) points out that there are many credible definitions of fairness, which can have dramatic impacts on the “fairest” solution for a given problem. In order to prevent self-interested “gaming” of the system, he recommends that participants in a negotiation agree in advance on an objective criterion of fairness. This is similar to the concept of the “veil of ignorance” central to Rawls’ Theory of Justice (1971) because, without knowledge of how it affects one’s own well-being, a person will likely choose what they truly believe to be fair. An alternative, more pragmatic, view of the same idea lies in the game theoretic heuristic that the best way to avoid “gaming” is to make the game too complex for players to discern what will improve their outcome, which may be true for the design for some large multi-stakeholder systems. “Macro” framing can also be an influencing factor in the creation of metapreferences on non-functional attributes in the design space for decision makers, such as a favoring of passively robust systems over actively changeable systems despite all-else-being-equal.

Micro framing. In contrast to macro framing, micro framing resides *within* the problem formulation, in the way information is presented and tasks are performed. The most prominent results in this field include bounded rationality (Simon, 1957) and Prospect Theory (Kahneman and Tversky, 2000), which delineate how humans may attempt to act rationally but do not succeed. Bounded rationality refers to the inability of humans to accurately analyze complex problems and find optimal solutions, instead relying on heuristics to reduce the cost of deliberation. Prospect Theory is an empirically derived theory describing the nature of many common deviations from axiomatic rationality. It states that people make decisions by comparing outcomes to a specified reference point. Outcomes are judged as differences from the reference point and are therefore perceived as “gains” or “losses.” Reference points are created from available information and are reinforced by anchoring, the observed bias that humans display towards information they are shown first, regardless of its ultimate relevance. Changing a reference point, once established, usually requires a deliberate effort (Tversky and Kahneman, 1974). Perceived value around the reference point is asymmetric, resulting in a higher impact of losses over gains as pictured in Figure 20. It has also been found that decision

making in the losses domain is more stressful and more likely to lead to irrational or regretted behavior (Gelfand et al., 2004).

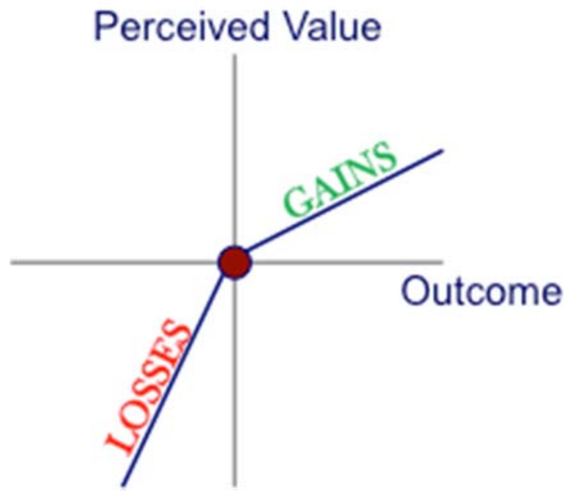


Figure 20 Perceived value around a reference point, according to Prospect Theory

Another common bias covered in Kahneman and Tversky's work is the availability bias, which describes the human bias towards information that is accessible (Tversky and Kahneman, 1974). In this way, information that is provided or readily recalled is implicitly assumed to be more important than hidden or forgotten information. Other biases they cover include insensitivity to probability, misconception of chance, and improper grasps of regression and representativeness.

The phrase *framing effect* is often used in the context of micro framing to denote an observable change in behavior derived only from changes in framing, usually with regards to whether the outcome is characterized as a gain or a loss. For example, switches between positive and negative (gains / losses) phrasing have resulted in dramatic changes in decisions, with people tending to strongly avoid losses over seeking gains (for examples and analysis, see Tversky and Kahneman, 1981; Levin et al., 1998). Additionally, the observed bias toward certainty has led people to be largely characterized as risk averse for gains, preferring a certain gain to a higher expectation uncertain gain, and risk seeking for losses, preferring a chance at no loss to a guaranteed loss.

Other topics in "micro" framing include the effects of detailed deliberation and expert opinion. Some research has suggested that extensive consideration of preferences can lead to behavior that deviates from expert opinion and leads to decreased satisfaction in decision outcomes (Wilson and Schooler, 1991). Excessive time spent developing a numerical value model, often without seeing the impacts immediately, effectively codifies the *estimation* as a *truth* that must be followed when more satisfaction would be gained by allowing future changes in response to emergent insight. This is a strong argument against overtaxing decision makers, and relates to the theory that the expertise of 'experts' is in fact dependent on a stable frame for them to leverage (Shanteau, 1992).

Finally, the concept of two-path information processing, a theory originally developed in the 1980s by the Elaboration Likelihood Model (Petty and Cacioppo, 1986) and the Heuristic-Systematic model (Chaiken et al., 1989) and recently popularized by Kahneman (2011), outlines two main ways in which humans perceive information and make decisions: heuristically and systematically (in ELM parlance, peripherally and centrally). Heuristic thinking is fast, developed over time and through intuition, allowing people to rapidly assimilate new information that they can fit into an existing mental frame. Systematic thinking is the more in-depth, analytical thought that promotes new learning but requires more effort on the part of the decision maker. The framing of a problem has an impact on which path a decision maker uses, depending largely on how familiar the situation is to them.

Framing in MSTSE. Prior research by the authors was specifically geared toward improving the micro framing of TSE/MSTSE visualizations in order to accurately represent the complexity of the multi-stakeholder problem and promote positive negotiation tactics (Fitzgerald and Ross, 2015). The benefit-cost tradespace scatterplot was predicted, based on the principles of negotiation theory and Prospect Theory, to emphasize the Pareto front as a reference point, thereby potentially miscategorizing some alternative as losses (relative to the front) when they are actually gains (relative to the *best alternative to a negotiated agreement*, or BATNA). Experimental evidence has lent credence to this theory.

However, this theory addressed only a fraction of the complete MSTSE process: micro framing in the analysis phase. To this point, very little consideration has been given to controlling macro framing or the framing of MSTSE problem formulation or modeling activities. Additionally, the benefit-cost scatterplot, though the most prominent tradespace visualization, is far from the only type of exploratory aid used in tradespace exploration. The need for further investigation of all aspects of framing in MSTSE is necessary in order for it to proceed as a viable means of engaging stakeholders in complex systems engineering negotiations.

Macro and micro framing can have a "weakest link" relationship in a negotiation, by which a framing trap in one may pull down the other. For example, if a stakeholder approaches MSTSE with a macro frame that is highly confrontational and individualistic, they will likely favor a micro frame, in the form of a particular visualization for example, that matches their outlook. Alternatively, if only individualistic visualizations are available, a stakeholder's macro frame may be slowly pushed into a similar aggressive, value-claiming mindset in order to reduce cognitive dissonance with their tools. For example, imagine a stakeholder with access only to a list of individually-Pareto-efficient alternatives. Naturally, he will be forced to engage with other stakeholders in the negotiation from the macro framing perspective of "I need one of these designs" rather than "we should find a mutually beneficial design" because he simply does not have the micro frame necessary to identify which alternatives on his list are agreeable to other people and therefore mutually beneficial. This defeats the central purpose of productive negotiation and because of it, management of framing must be continuous, extending from problem formulation all the way through analysis.

Framing Activities and Visualizations

Using the generic three-step outline of a TSE procedure, the following sections will detail recommendations for controlling the framing of common TSE activities in order to support a successful MSTSE application. The success criterion of MSTSE is the ability to find and identify mutually beneficial alternatives, if they exist. To do that, the macro framing of the problem should be aligned with the tenets of principled negotiation as much as possible and the micro framing must accurately represent the value of the different alternatives. The recommendations included here are not intended to be exhaustive but rather instructive advice for potential adopters of MSTSE, based on the combined insights of literature in framing and negotiation. Following these recommendations should improve the communication of preferences and needs between negotiators (a skill not developed or supported by classic TSE) and the value assessment of the alternatives by each negotiator (which is a different, more complex task than in classic TSE). This improves the MSTSE procedure by reducing the likelihood of key failure modes at both the inter-stakeholder and stakeholder-data interfaces, limiting opportunities for negotiation breakdown driven by social conflict or misattribution of value.

Problem Formulation

Problem formulation has a large impact on the resulting direction of a tradespace analysis. It defines the scope of the system to be analyzed, what factors are (and are not) under designer control, and the sources of value that are sought by the stakeholders. Unsurprisingly, the predominant impact of framing in this stage is likely to come from macro framing as the beliefs, perspectives, assumptions, and sometimes biases of the participants work their way into the problem. To address this challenge, communication becomes paramount: explicitly capturing some of the macro frames with which stakeholders and/or analysts are approaching the problem can allow for the identification and mitigation of potential future barriers to agreement before they become negotiation impasses.

Capture macro frames. Note that the objective of these efforts is not to *change* the macro frames with which stakeholders approach the problem, but to capture what they are. Practically, macro frames are developed by a lifetime of experience and opinion, and are difficult to change. More fundamentally, since MSTSE is positioned as a *prescriptive* rather than *normative* analysis technique, it is inappropriate to suggest that one macro frame is the "correct" frame to use (a normative argument). Rather, we are interested in knowing the macro frames favored by each stakeholder so that when *they* attempt to make a normative argument we can understand the frame leading them to make that argument and, hopefully, communicate it effectively to other stakeholders who do not share that frame. This is intended to prevent incidents of the stakeholders "talking past" each other by assuming others share their underlying assumptions.

Some useful frames to consider are:

- Purpose for MSTSE (e.g., to explore and learn about the opportunity vs. to make a funding decision)
- Relative desire for low-cost vs. high-benefit systems
- Relative desire for passively robust vs. actively flexible systems

Record key elements of problem structure. This activity is already a main component of problem formulation for TSE, which requires explicit accounting of the factors impacting the system and their assignment as variables in the tradespace: design variables, context variables, or performance attributes. However, the multi-stakeholder problem has additional structural elements on top of those from single-stakeholder tradespaces that can impact the best micro frames to use in later phases of MSTSE. Explicitly noting these elements during problem formulation can improve later analysis, as certain analysis types can become more or less relevant depending on these key features. For example, if some attributes of interest to the stakeholders are divisible at-will (e.g. manufacturing costs, which can be split between stakeholders as desired), these can be leveraged by additional analysis later by customizing or sub-optimizing a given alternative. On the other hand, negative pre-existing relationships between the stakeholders may limit the effectiveness of some types of exploration, particularly those that involve directly comparing desired alternatives. Some of the structural elements worth recording include:

- Divisible attributes
- Relationships between stakeholders (personal, professional, etc.)
- Tradespace completeness – could more alternatives be added?
- Constituencies – do the stakeholders represent other people?
- Schedule – how much time is available for the stakeholders to interact?

Determine each stakeholder's BATNA. This is arguably one of the "key elements of problem structure" from the previous point, but is critical enough to merit its own description. The BATNA (best alternative to a negotiated agreement) is, essentially, what each stakeholder will do *on his own* if no agreement can be reached with the other stakeholders. This is an important reference point with respect to the value of any of the design alternatives under consideration as it defines the border between gains and losses. Failure to define and then leverage the BATNA during exploration reduces the situational awareness of the stakeholders.

In some cases, the BATNA will be readily apparent, particularly if the stakeholder(s) have *no* viable alternatives to a negotiated agreement. However, in general this task requires careful thought and consideration just like the rest of problem formulation. It can help to consider a variety of "types" of BATNA, in order to prompt brainstorming in multiple areas. Common BATNAs include the following:

- **Do-nothing** – if the MSTSE is strictly exploratory, inaction is likely the course of action should no agreement to proceed be made. Doing nothing typically carries zero cost and zero benefit.
- **Existing system** – for design tasks intended to improve or replace an existing system, the do-nothing alternative actually entails using the current system. This type of BATNA is one that commonly drives differences in stakeholders' bargaining leverage, as some stakeholders may be much better off with the current system than others.
- **Build preferred alternative alone** – some projects seek agreement between multiple stakeholders to reduce the cost borne by each individual. If a stakeholder is capable of affording some or all of the alternatives by themselves, those alternatives become viable BATNAs (though at a higher cost than if they could agree to share one).
- **Other opportunity** – resources that are expended on the alternatives in the tradespace represent an opportunity cost in that they cannot then be spent on other projects, which may be more valuable. This type of BATNA is the most difficult to capture, as the number of other opportunities is potentially limitless, but this fact is true for all design tasks. Usually a small number of known viable or attractive opportunities can be considered without fear of missing drastically better choices.

Identifying the best alternative in each of these categories and then assigning the best of those as the BATNA is an effective way of breaking down the problem. Sometimes it may be difficult to assess which of these choices is the "best" (and thus, the BATNA) at this point, because the evaluative model has yet to be created, particularly for the "build alone" choice. In that case, preserving the list of potential BATNAs and then choosing one after modeling but before exploration is feasible.

MODELING / EVALUATION

Engaging in the modeling of the system after completing a thorough problem formulation seems at first glance to be trivial: simply a matter of taking the defined design vectors and finding the right equations to calculate the desired performance attributes, subject to any influencing contextual parameters. However, the modeling task itself can also propagate cooperative versus individualistic framing implicitly into the exploration phase. When multiple stakeholders will be conducting the exploration, it is important to make sure that the modeling is satisfactory to all of them, which requires some additional management.

Joint Fact Finding (JFF). Joint Fact Finding (Ozawa, 1991; Ehrmann and Stinson, 1999) is a valuable use of time in order to build trust in the data that exploration will be based on. It is difficult to reach consensus on a design if some stakeholders disagree with the models being used to evaluate it, making uncoordinated multi-person modeling activities a threat to productive negotiation. JFF seeks to establish credible and objective data, one of the foundations of principled negotiation (Fisher, Ury, and Patton, 1991), to use as the foundation for evaluation of alternatives and discussion of their relative merits. If possible, all efforts should be made to convene stakeholders prior to actual exploration in order to perform JFF in support of the modeling task. JFF also helps to establish a macro frame of cooperation *before*

engaging in the negotiation itself, which can help preserve positive, mutually-beneficial bargaining in the face of any naturally developing competitiveness.

Private Information. Not all models can be developed through JFF. If a stakeholder already possesses a model for a piece of the larger system, reusing that model can save time and effort. If they are willing to share that model (both how it works and its results) with the rest of the stakeholders as a part of a larger JFF effort, that is a valuable step in building rapport, in accordance with the principle of Full, Open, and Truthful Exchange (Raiffa, 2002). Some stakeholders may be reluctant to share models, but should be encouraged to do so for the above reasons. However, some models' inner workings may depend on proprietary or classified information that the stakeholder is unable to share. In the case of a stakeholder unwilling or unable to reveal their models, two approaches can be taken: the existing model can either be ignored in favor of a newly-created JFF model (if possible) or "black-boxed" so that other stakeholders can only see its outputs. A black-boxed model can be fully effective if its outputs only impact the value proposition of the stakeholder who owns it. If not, other stakeholders will need to trust that the model is accurate. If a public - but presumably lower fidelity - model is available, it can be used to help validate the black-boxed model and build trust.

EXPLORATION / ANALYSIS

Entering the exploration phase, the dominant framing concern shifts to micro framing: the actions the participating stakeholders are asked to perform and the way the data generated by the previous steps is presented. Macro framing still has a role to play in exploration however, specifically when weighing specific alternatives as potential final agreements.

Emphasize the BATNA. For a proper valuation of the designs in the tradespace, they must be valuated against the BATNA as a reference point. This provides the necessary perspective for determining the value of a design *as a multi-stakeholder agreement* rather than the typical, less-contextualized evaluations *in a vacuum* or *relative to other designs* commonly used in classic TSE activities. Taking classic TSE visualizations and intelligently incorporating a prominent indicator of the BATNA is a functional way of improving negotiation behavior, as demonstrated by the negotiation tradespace in Figure 21: the use of which was shown via controlled experiment to improve gains/losses framing with a more accurate reference point (Fitzgerald and Ross, 2015). Views designed to compare alternatives should always include the BATNA as a "sticky" alternative, even in simple implementations such as tables of performance data.

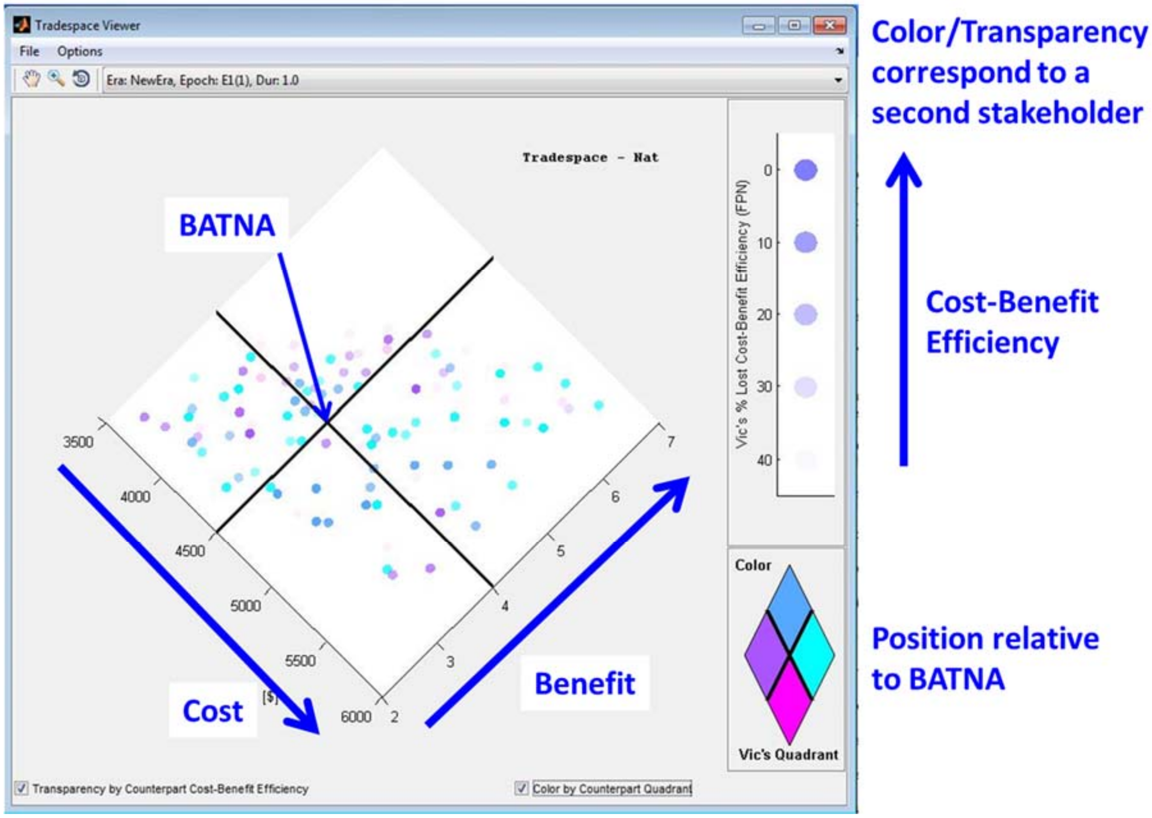


Figure 21 Negotiation tradespace used in MSTSE experiment (Fitzgerald and Ross, 2015) with key features

Limit strictly-individual analysis. Activities should incorporate the value statements of multiple stakeholders as much as possible in order to consistently keep each participant aware of the "group" aspect of the negotiation problem. This can prevent fixation on alternatives that are very good for one stakeholder but not for others. In the BATNA-centric tradespace, color and transparency accounted for the value of other stakeholders, and the resulting negotiations saw fewer exhaustive search patterns in favor of more direct paths to mutually-valuable solutions. If the participating stakeholders want to utilize a particular analysis of the tradespace using their own value, it should be replicated for other stakeholders and shown together. For example, the benefit-cost efficient solutions on the Pareto front are highly desirable for a given stakeholder, but should be calculated and presented relative to the Pareto fronts of the other stakeholders. This can be accomplished in multiple ways, including the use of Venn diagrams to illustrate overlap between specific stakeholders' preferred alternatives and gridmaps to show the relative sizes of the regions of agreement for all stakeholders (Figure 22). These, and other, visualizations are currently the subject of ongoing research.

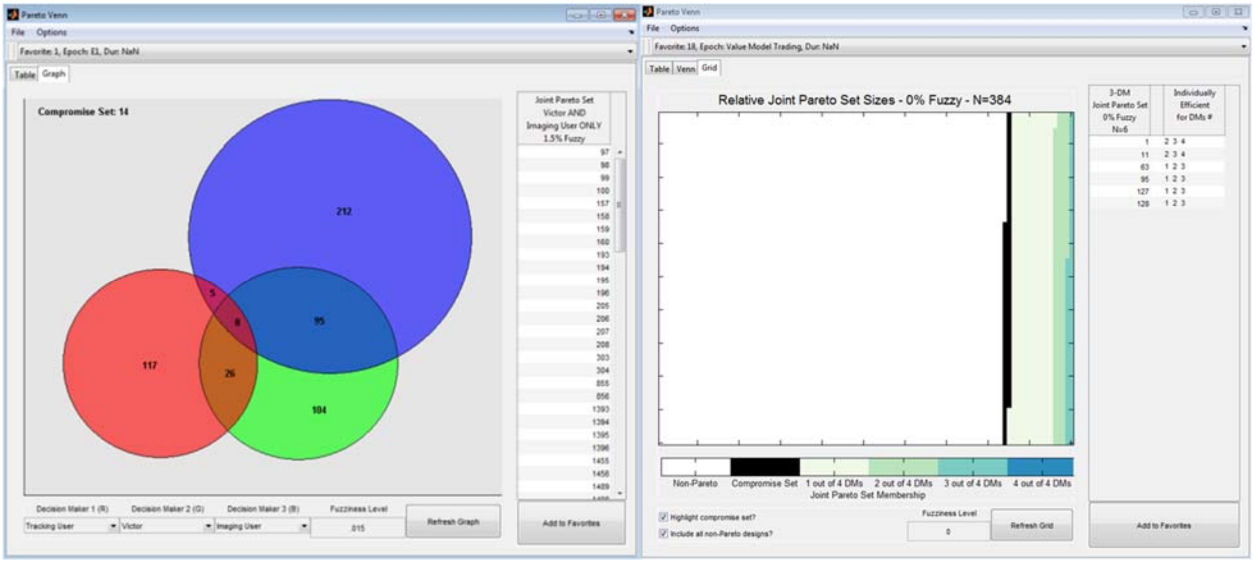


Figure 22 Example Venn diagram and Gridmap for multi-stakeholder Pareto front analysis

Analyze relationships. The relationships between stakeholders in the value domain is a component of the multi-stakeholder tradespace that is not present in classic TSE, but is just as important as the evaluation of the alternatives directly. These relationships, whether or not they are analyzed, will affect the ways stakeholders interact and the designs that they might agree on; thus explicitly considering them is a powerful means of understanding the dynamics at play in the negotiation. Stakeholder relationships in the value domain can be quantified through the correlation of their value metrics, commonly done at the holistic level (e.g. the correlation between Stakeholder A and Stakeholder B using their respective cost-benefit efficiencies) and displayed in a heatmap for all stakeholders at once. This view can visually highlight groups of stakeholders that could form a promising coalition of shared interests, which can be a useful simplification of a many-party negotiation; in Figure 23, separate three-stakeholder and two-stakeholder coalitions with internal correlation greater than approximately 0.6 are apparent in the blocks of green, with an average of approximately 0 correlation between the two coalitions. Additionally, explicitly showing positive correlations indicative of shared interests can be a useful reminder of the potential for mutual gains for stakeholders caught up in a distributive negotiation fallacy or fixated on individually-optimal alternatives.

Correlations can also be displayed on an interest-by-interest basis (e.g. the impact on the correlation of A and B's utility functions caused by A's preference on a specific value metric). The resulting correlation data is combinatorically larger than at the holistic level but can be segmented to provide an intuitive breakdown of how one stakeholder relates to all of the others. This can be used to identify key "free" attributes that do not need to be traded between stakeholders and "pain points" that drive the differences in the value statements for each stakeholder. In Figure 24, the orange-coalition stakeholder Victor's attributes are displayed on the y-axis and it becomes clear that his second attribute ("NumTargetBoxes") is driving a considerable part of both his alignment with his own potential coalition and disalignment with the other stakeholders.

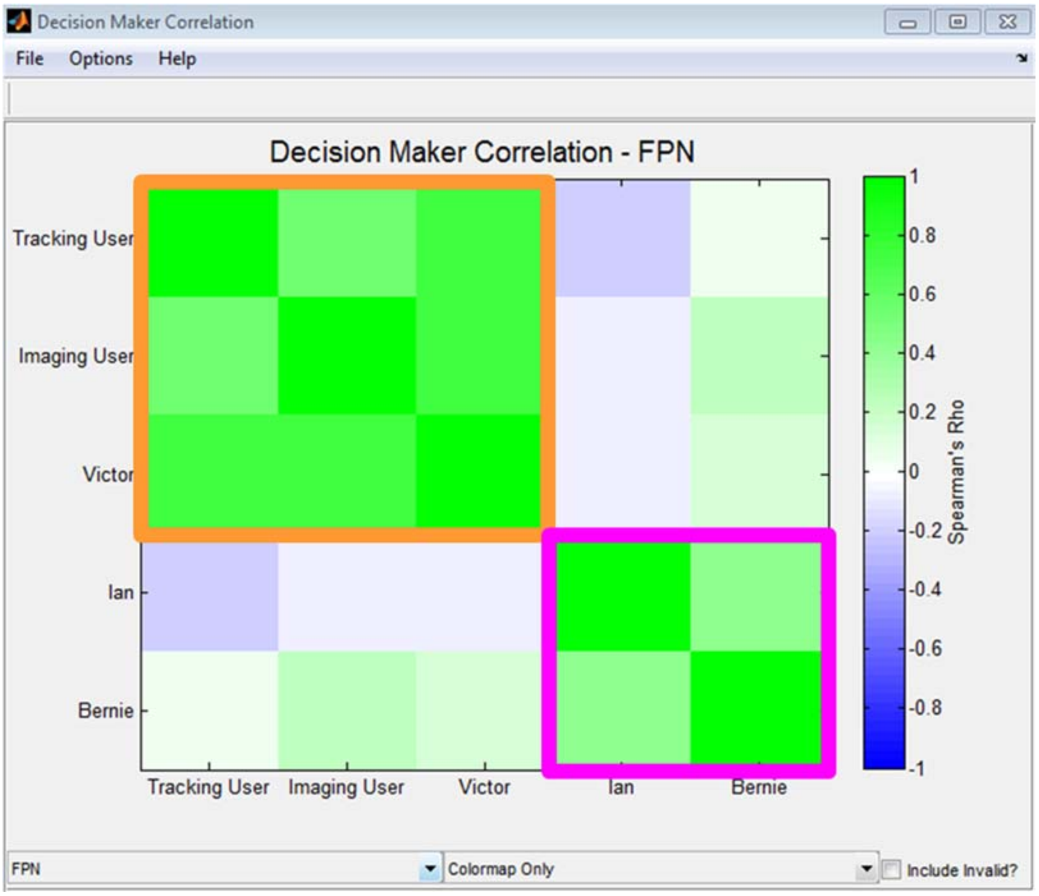


Figure 23 Example Stakeholder-Stakeholder correlation interface for five stakeholders – annotated to highlight two emergent coalitions with high correlation (orange and magenta boxes)

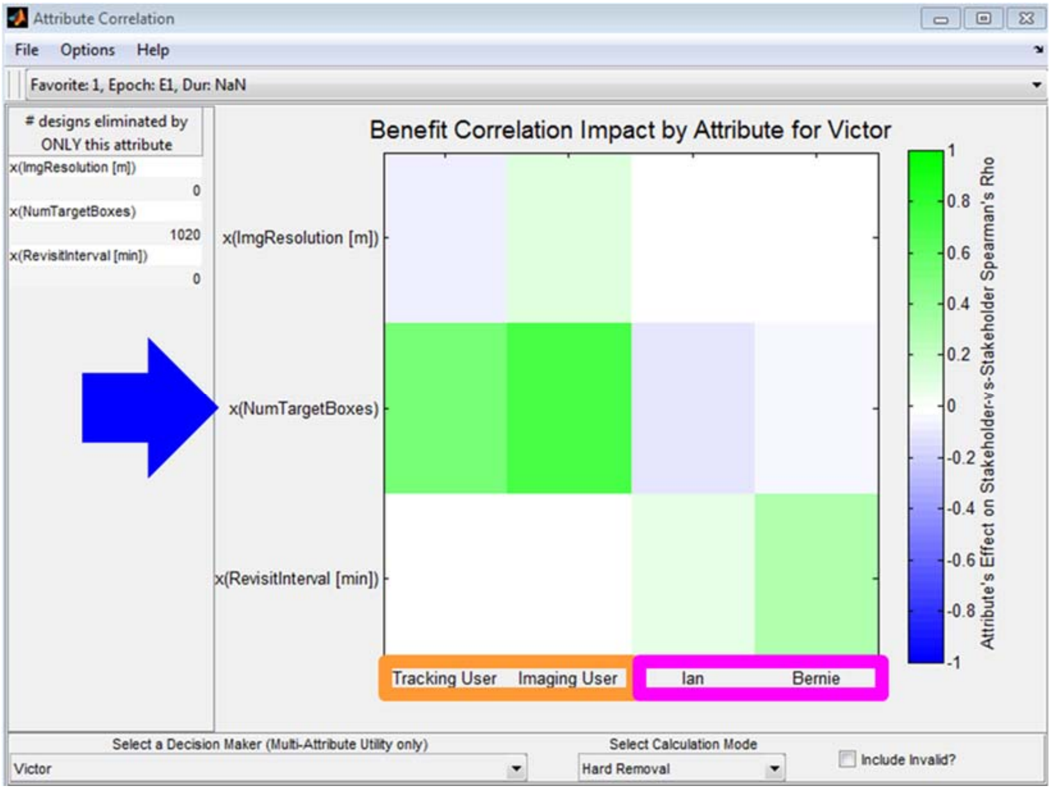


Figure 24 Example Stakeholder-Interest correlation interface – annotated to highlight the attribute most responsible for bringing the orange coalition together and separating them from magenta

Allow stakeholders to change their mind. Negotiation in MSTSE exposes each stakeholder to large amounts of information that they may not have previously known, particularly the preferences of other stakeholders which are not present in classic TSE. New information can change subjective assessments of value (Curhan et al., 2004) and invalidate parts of the original problem formulation. Stakeholders should be encouraged to critically reassess their value statements during the negotiation. Tweaking the value functions to more closely align with a "new" reality can be performed during a session in order to accelerate the iterative design loop (see Figure 25 for example). Additionally, if the value function updates are convergent in a manner leveraged by other consensus-building techniques such as the Delphi method (Golkar and Crawley, 2014), these live updates have the potential to open up new regions of mutual value in the tradespace.

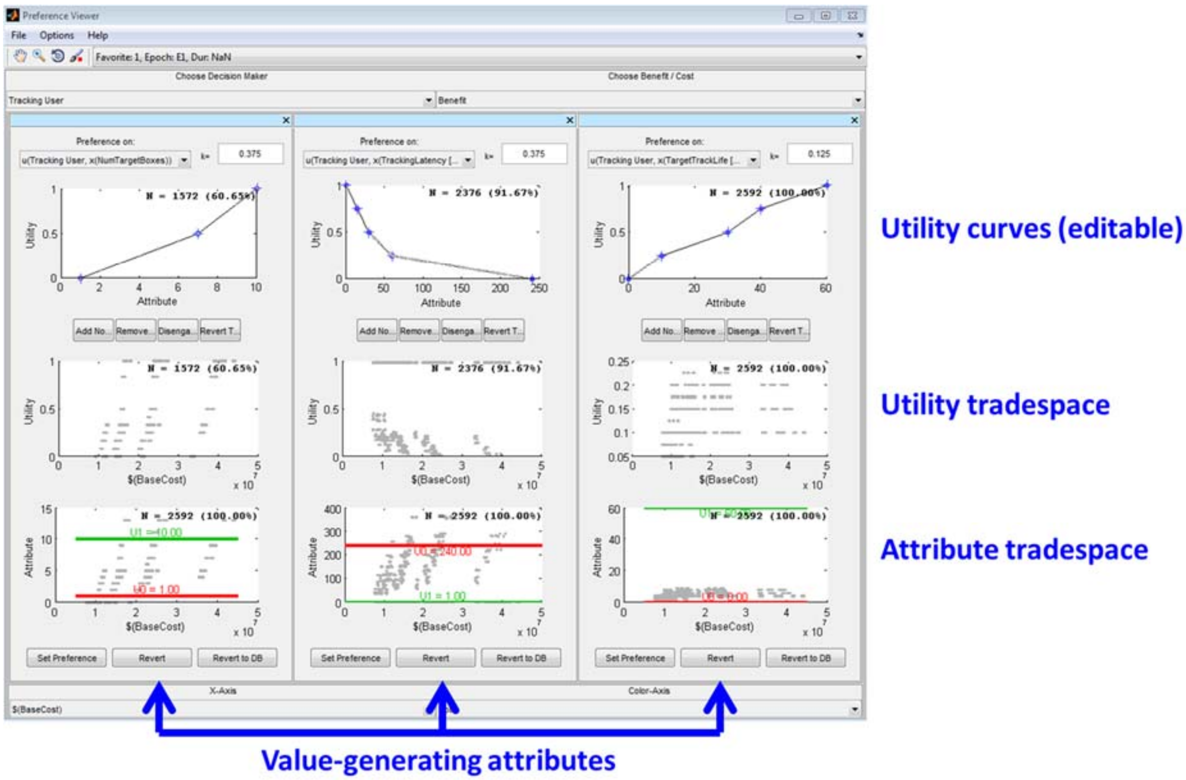


Figure 25 Example interface for live editing of value functions, synced to other visualizations

Refer back to macro frames. When discussing individual alternatives, effort should be made to refer back to the macro frames of each stakeholder. When a stakeholder refers to a design with a subjective assessment like "good", the first question should always be "Why?". Each stakeholder wants a "good" design, but each has different criteria for what is "good" that includes not only their reported value function but also the macro frames with which they choose to make decisions. For example, if Stakeholder A recommends an alternative as "good" on the grounds that it has high benefit for all parties, Stakeholder B can make a more intelligent counteroffer with less chance of sparking a debate over the definition of "good" if it is clear to all parties that he prefers low-cost, high-efficiency solutions over strictly high-benefit solutions.

Framing Multi-Stakeholder TSE: Discussion and Conclusion

TSE is a continually developing design paradigm, and MSTSE is an even younger offshoot of the main research branch. Considerable work is still needed to flesh out the similarities and differences inherent in exploring a tradespace with one stakeholder versus multiple stakeholders, particularly in the realm of implementation. Framing has the potential to elevate or sabotage group analysis depending on its suitability. This work is an initial attempt to identify framing activities necessary for MSTSE, and to provide recommendations for how to conduct them to greatest effect. Importantly, these framing activities span problem formulation, modeling, and exploration and include both macro framing and micro framing concerns.

This section has addressed the framing of MSTSE with active stakeholder participation from problem formulation through exploration – as opposed to "informal" MSTSE relying entirely on engineers and/or analysts, which was mentioned briefly when introducing the evolution of the topic. Given the many constraints on most stakeholders' time, informal MSTSE will likely remain a practical alternative for developing insight into the dynamics and relationships that define multi-stakeholder problems. However, lack of stakeholder participation imposes some limitations on the types of activities that can be performed effectively. Stakeholder value models and BATNAs will need to be estimated and can't be modified during exploration (as a stakeholder can't "change their mind" without participating). Additionally, some tasks will revert to their standard TSE forms, as Joint Fact Finding is not possible and stakeholders will not be available to discuss macro frames. Table 2 presents a short summary of recommendations in this section and the modifications necessary for their adoption in informal MSTSE.

Table 2 Summary of recommendations, with modifications for informal MSTSE

Phase	Recommendation	Informal MSTSE
Problem Formulation	Capture macro frames	All of these apply except for capturing macro frames of other stakeholders. Make best estimates for stakeholders' BATNAs and value models.
	Create many alternatives	
	Record key elements of problem structure	
	Determine each stakeholder's BATNA	
Modeling / Evaluation	Joint Fact Finding	Treat modeling as normal TSE
	Private information	
Exploration / Analysis	Emphasize the BATNA	Continue to use BATNA-centric visualizations and analyze relationships, but limit activities related to changing stakeholder value models without their participation.
	Limit strictly individual analysis	
	Analyze relationships	
	Allow stakeholders to change their mind	
	Refer back to macro frames	

Originally, MSTSE was envisioned to leverage the TSE framework in order to capture insights from the data related to the multi-stakeholder dynamics of the problem and find better negotiated solutions. Explicitly managing the framing aspects of MSTSE can serve to enable this goal by reducing opportunities for the social breakdown of negotiation caused by poor communication or degenerate bargaining tactics, which can occur at the stakeholder-stakeholder level or at the stakeholder-data level. The framing elements called out in this section represent a first pass at collecting some of the most important features of the MSTSE technique; future research will seek to expand on this list and provide more actionable recommendations for practitioners. Additionally, future research will expand the validation efforts of previous experimental results (Fitzgerald and Ross, 2015) by incorporating the insights of case studies, both by considering the impact that the framing of issues has had on negotiations and by testing the ability of the MSTSE framework to identify "good" and "fair" solutions under a variety of potential macro and micro frames held by the participants.

MODEL CURATION

The need for consideration of curation in model-centric engineering stems from the significant increase in models and model-related assets (libraries of data sources, techniques, etc.). Figure 26 illustrates the many decisions one would need to make in using models on a complex systems endeavor, such as an airport collaborative decision making system. This brings into question who owns such assets at an enterprise level, and how these are managed and used/reused over time.

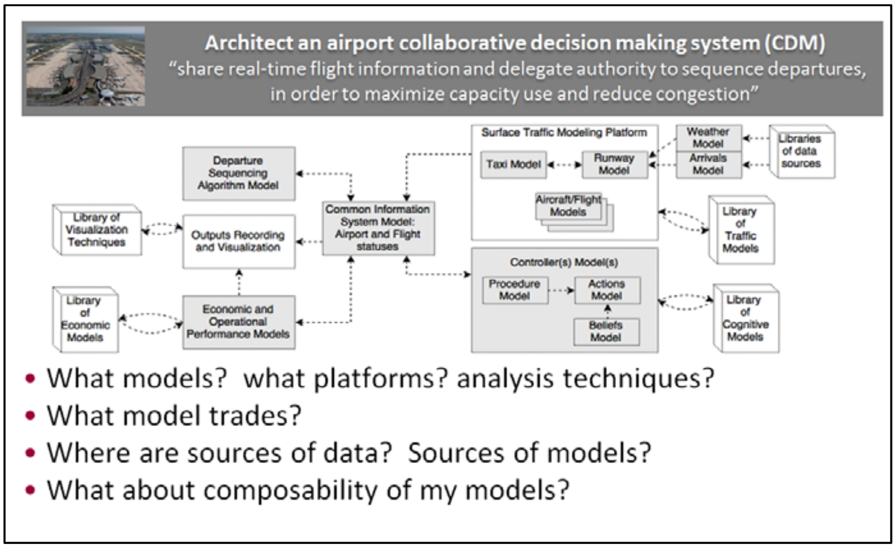


Figure 26 Many Decisions in Using Model-Centric Environments

Curation in model-centric engineering is a relatively new topic, but the science of curation in other fields (museum curation, digital curation, content curation, etc.) provides useful constructs and approaches to inform curation in the systems engineering context. During this phase of the research, knowledge gathering from other fields has been ongoing, alone with discussions with experts in the field of systems to define needs and applicability of general curation approaches.

BACKGROUND

Model curation was introduced as a SERC topic in the IMCSE phase 3 research report (Rhodes and Ross, 2016). In this subsection, we repeat some content as background for the reader.

The January 2015 IMCSE workshop participants cited model curation as an important topic for investigation in evolving model-centric engineering (Rhodes and Ross, 2015). Rouse (2015) stresses that the wealth of existing models is often not used because of a lack of knowledge of these resources and the difficulty in accessing them. As engineering practice becomes increasingly model-centric, models are valuable assets for designing and evolving systems. The need for model curation accordingly becomes a necessary functional role in organizations. This

is a relatively new idea in the systems community, though progress has been made on some of the activities involved in curation. Maturing an approach for model curation in the systems engineering field can leverage the work of other related curation practices. Digital curation experts from the U.K.'s Digital Curation Centre previously noted "As scholarly research and scientific study becomes increasingly driven by the analysis of data, long term access to these data is crucial in enabling the verification of scientific discovery and to providing a data platform for future research" (Rusbridge, et al., 2005). As digital curation is closely related to modeling curation, there is much to borrowed and adapted from this practice. Practices on collaboratively developed model repositories and their management provide additional insights for model curation (EMBL, 2015). Another related area is social content curation, focusing on collaborative sharing of Web content organized around one or more particular themes or topics. "Social curation can be defined as the human process of remixing social media content for the purpose of further consumption" (Duh et al., 2012). It is viewed as a complement to traditional data exploitation methods such as algorithmic search and aggregation. Curators will need to know about a number of things including model ontologies, model meta-data, latest modeling techniques and classes of models, policies on data rights, code of ethics, and others.

Effective model curation necessitates clarity across the systems community in characterizing and handling models. It requires formalizing knowledge of models and determining a distinctive set of model characteristics (purpose, input/output types, logic, assumption types, model incompatibilities, etc.). Various types of models have been enumerated by the systems engineering community, but there appears to be insufficient attention given to model purpose itself. In the field of informatics, McBurney (2009) proposed nine model purposes, from understanding, predicting or controlling natural reality (e.g. Newton's laws) to playing and enabling the exercise of human intelligence, ingenuity and creativity, in developing and exploring the model (e.g. so-called serious games). Zacharias et al. (2008) point out that individual, organizational and societal models, do not predict exactly what humans will do, as individuals or in groups, but rather help forecast a range of potential action outcomes, draw attention to potential unintended consequences, and highlight variables that are overlooked in a particular situation.

Accordingly, model purposes include: to analyze fragmented information and develop courses of action based on the likelihood of desired outcomes; to train personnel, simulating the environment, dynamics and providing performance feedback; and to design and evaluate a technical system, predict its performance and make decisions based on cost-benefit tradeoffs. Recent work on hybrid modeling and multi-scale modeling point towards the usefulness of using not a single but a set of models to study a complex system (La Tour and Hastings, 2015; Mathieu et al., 2007; Zulkepi, et al., 2012). Curators of model-centric environments, similar to a museum curator, could help select the most appropriate set of assets given a purpose. Bankes (1993) opposes consolidative computational modeling of deterministic systems (models as surrogates for real systems) to exploratory modeling of systems plagued with uncertainty and unknown unknowns (models as means of testing hypotheses and exploring ranges of possible outcomes). It is argued that exploratory modeling can only produce useful results through a constellation of alternative models. By using multiple simple models, complexity is exported

outside the models to the ensemble of model outcomes, from which modelers and stakeholders must make sense. Ross et al. (2015) demonstrate the importance of model tradeoffs and choice in tradespace exploration. Selection of a useful set of models given a specific system and modeling purpose requires specialized skills; that could be an area of expertise of a curator. And, effective model curation practices could ensure availability and access to a validated set of models with associated pedigrees (Reymondet et al., 2016).

As the model-centric environments become increasingly complex and critically important, there is a need to more strategically lead and manage them. Under the premise that model-centric environments of the future will necessitate specialized leadership and competencies, a new leadership role for curation has been investigated. The curation function would set and administer model-related policies and practices, and ensure models and related documents are authenticated, preserved, classified and organized accordingly with model metadata standards. The curator may own the data management for models and related information, or oversee the ownership by other individuals or organization. As needed, a curator would meet with individuals and teams, who will create, use and re-use digital assets, helping to determine a useful classification of both individual models and sets of models. At the organization level, the curator may organize training and special projects. Empirical knowledge gathering has investigated the challenges and needs, and investigated the potential roles and responsibilities for this curation role.

As engineering continues to evolve to a model-centric paradigm, there are many challenges and considerations. There is a need to better understand *where* the role of the human (versus automation or “AI”) is essential in the effective management and utilization of model environments. This includes the models (managed as organizational assets), supporting infrastructure, and the associated protocols and practices. Digitized legacy systems and new digital system models will provide the basis for designing and evolving systems into the future. This drives the criticality of models as assets and necessitates change in model-related policy and practices. Accordingly, there is an urgent need to mature a practice of “model curation” including a “model curator” functional role within engineering organization.

In the envisioned future, model-centric environments will be under the leadership, oversight and management of a curator function. This function (performed by a group of individuals) will have the objective of sustaining highest possible benefits and outcomes from the collective set of model assets and formal curation practices. Reymondet et al. (2016) investigates considerations for curation in the engineering of complex socio-technical systems. Extending from the various types of curation roles and activities of other fields, the model curator’s role is envisioned to include a number of major responsibilities and support of various staff.

The model curator (curation function) would set and administer model-related policies and practices. The curator would ensure models and related documents are authenticated, preserved, classified and organized accordingly with model metadata standards. The curator may own the data management for models and related information, or oversee the ownership by other individuals or organization. As needed, a curator would meet with individuals and

teams, who will create, use and re-use models, helping to determine a useful classification of both individual models and sets of models. At the organization level, the curator may organize training and special projects related to model-based engineering. The goal will be to develop a comprehensive roles and responsibilities description, and gather findings that may support the development of a model-centric environment self-assessment.

DRIVING FACTORS

IMCSE research suggests a number of driving factors for model curation and a curation leadership role at the enterprise level. Although reuse of models can have benefits, the reality is that legacy models are not widely used beyond their original purpose. Accordingly, modeling efforts are often duplicated across programs. The lack of access to models, mistrust of models, and perception of legitimacy of models are all barriers in model reuse.

In many enterprises modeling competency is distributed across individuals and organizations, rather than leveraged at the enterprise level. A lack of a centralized leadership authority means that models are owned and managed only at the local level. Since model expertise is largely resident in individuals, the ability to select and compose sets of models is typically limited to the original use. Programs that lack model experts accordingly do not benefit from the collected wisdom of the enterprise.

The question arises, could a curation function at the enterprise level (Figure 27) lead to more effective model-centric environments that enable more effective use of models and model-related assets?

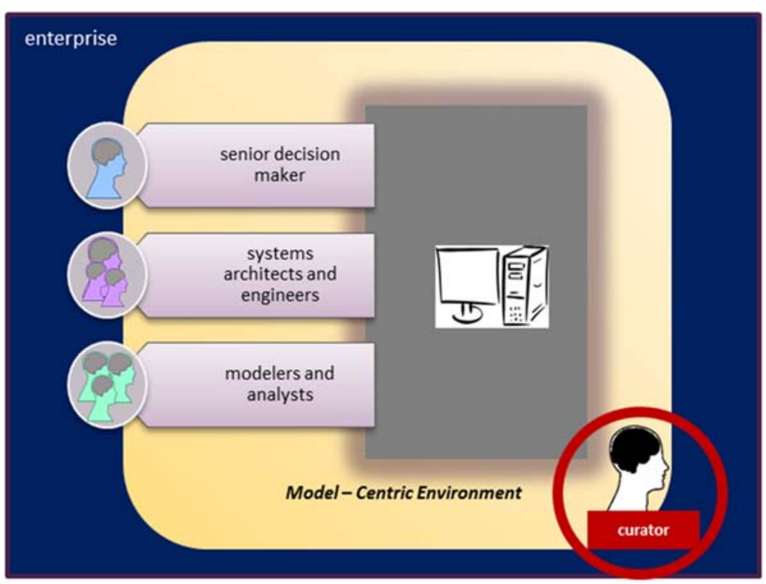


Figure 27 Enterprise-level curation of model-centric environments

KNOWLEDGE AREAS

Successful model curation efforts will require knowledge of myriad aspects of models and model practices. Figure 28 highlights six areas, which are briefly described below.

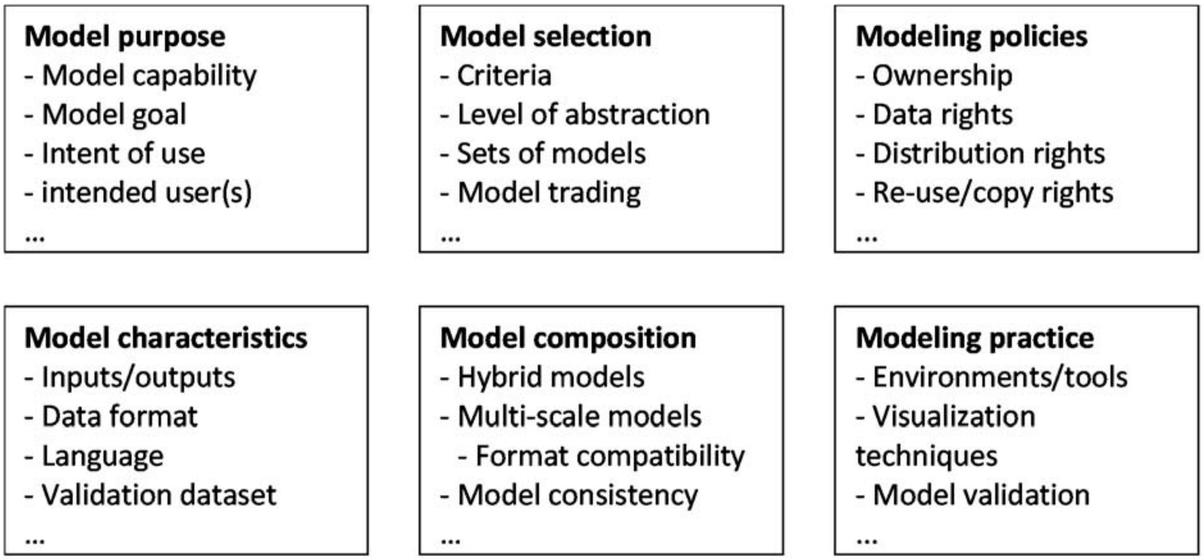


Figure 28 Knowledge Areas for Model Curation

Model Purpose. The model purpose specifies its intended use, and sets the boundary for model depth and breadth. The capability of the model as related to its overall project goal provides important descriptive information. Models are valuable within the context of the project, so detailed information is needed on intended use of the model within this context and the intended users. A shared understanding of model purposes is necessary for appropriate choice of an adequate set of models.

Model Characteristics. Models can be characterized in many ways. The rich set of information includes information about inputs and outputs, data formats, languages used, uncertainty, assumptions, and other metadata an organization deems as valuable. Another essential characterization is the validation approach and dataset, which is key information that drives confidence in a model. Model characteristics provide additional elaboration to the metadata that is needed to enable case-specific selection of models.

Model Selection

Models need to be carefully selected with appropriate criteria for the project and its context, as well as practical criteria (organizational standards, availability of model data, etc.) Traditional, often computational, methods for systems engineering are inadequate for a complex system such as a sociotechnical system, which includes soft and hard problems, a very large number of technical components and autonomous actors, and emergent behavior created by soft-hard, social-technical, or non-linear interactions. Using a set of models enables insight into various aspects of system complexity. In contrast with consolidative computational modeling of

deterministic systems (models as surrogates for real systems), exploratory modeling of systems involves greater uncertainty and unknown unknowns (models are useful in exploration as means of testing hypotheses and exploring ranges of possible outcomes). However, it is argued that exploratory modeling can only produce useful results through a constellation of alternative models. Selecting one or several models for a complex case study requires breaking down the problem into scoped questions such that each question translates into a model purpose.

Model Composition. Models are often combined with other models, such as in hybrid models and multi-scale models. This necessitates having knowledge of the format compatibility, consistency of the models, and justification for how and why these are composable. For example, interconnecting two models, especially two computational models, raises issues such as incompatibilities in data types, in naming schemes, in logic mechanisms, in execution timing.

Modeling Policies. Formal models often become valuable assets of not only the project, but the larger organization. A key area of knowledge for curation is modeling policies, including legal aspects and organizational policies. Organizational policy may specify model ownership, for example, and specify guidance for sharing and reuse of models internally to the organization. General knowledge of intellectual property law is necessary for model curation, particularly to inform when and where to seek legal expertise.

Modeling Practices. Models are used in context of organizational and/or project modeling practices, whether formal or informal. Model development lifecycle practices (e.g., model verification, model validation) need to be well-understood for effective curation. Understanding of the existing model-centric environment, toolsets, and techniques is fundamental for curation. It is also necessary to continuously acquire information concerning emerging practice and enabling tools and techniques.

MODEL PEDIGREE

Model pedigree is not a new idea. First described as "model demographics" by Gass and Joel (1980) as model demographics, the term pedigree is subsequently used by Gass. A pedigree contains all of the information about a model, its origins and use over time. As described by Gass and Joel (1980), the purpose is to "enable the decision maker to determine the model's status with respect to past achievements, theoretical and methodological state of the art, and the expert advice that went into its development".

While model documentation is typically developed, the content of the pedigree may contain information not always included in engineering model documentation materials. Given the preliminary work in this phase of IMCSE research, there is a need to evolve a standard pedigree for use in the systems community.

MODEL CURATION ROLE

The DoD Digital Engineering Working Group published a set of SE Digital Engineering Fundamentals (2016) as guidance, which included the following:

> *The responsibility of planning and coordinating programs' use of models, simulations, tools, data, data rights, and the engineering environment belongs to the program manager; the performance of the actual may be delegated to the program systems engineer and other program staff as appropriate*

The question arises as to whether this practice will be sustainable given the envisioned future under the digital engineering paradigm. It is plausible that enterprises will need a specialized leadership role that will enable enterprise-level management and control of digital assets. During this phase of IMCSE research, some preliminary investigation included discussions with research stakeholders and experts in the field on the topic of model curation and possible roles.

At the enterprise level, it is envisioned that there may be an executive leadership role of a curator, similar to a museum level curator. Looking at the responsibilities of institutional level curators provides insights for a potential role and responsibilities:

An executive level curator for the model-centric enterprise may have the following role and responsibilities:

- Acts as the executive process-owner for model-centric environments
- Provides governance for the data/model repositories, data rights, IP
- Protects the model 'pedigree' and ensures pedigrees are maintained
- Guides selection of models (individual and collections) and modeling platforms
- Owns/manages model risk and opportunity at the enterprise level
- Negotiates borrowing and loan of model assets with other enterprises
- Possesses deep, current knowledge of models, model trades, composability...
- Convenes panels of program-level model leaders
- Orchestrates demonstration of model-based capabilities to support bid and proposals
- Conducts model capability assessments
- Determines and assesses modeling competency in the workforce
- Manages the accessioning and de-accessioning of the enterprise's model assets
- Provides assurance of cataloging and tracking of model assets

At the program or business area level, a model curator has responsibility at the local level. As an example, local level responsibilities may include:

- Selects and maintains the set of models for a specific program or laboratory purpose
- Originates/updates model pedigree for the program-owned model assets
- Plans and manages model version upgrades
- Works with model software developers on specialized needs
- Organizes training of new program staff
- Supports enterprise-level model assessments and activities
- Performs model tradeoffs and model software selections

Lexicon (Preliminary)

As the dialogue around model curation continues, it is helpful to begin building a lexicon. Selected terminology is shown in Table 3 in a preliminary lexicon is proposed for further discussion and maturation by the systems community.

Table 3 Model Curation Lexicon (selected examples)

Terminology	Definition
Model Accessioning	The formal process of accepting and recording a model as a collection object in the enterprise level model portfolio. Accessioning addresses the legal, IP and ethical issues in model acquisition and development.
Model Cataloging	The formal process of making a model available for use and tracking it throughout the model lifecycle.
Model Collection	The collection of model-based assets that is owned by an enterprise
Model Collection Object	A model or model-related object that is a unique asset in the enterprise's collection of model-based assets
Model Curator	A designated professional role entrusted with the ownership, tracking and use of model-based assets
Model Composition	The process of composing a set of models and model-related information that provides collective value beyond the individual models.
Model Composability	The characteristic of an interrelated set of models to be combined in accordance with given modeling formalisms.
Model De-accessioning	The formal process of removing a model as a collection object in the enterprise level model portfolio.
Model Metadata	Metadata means simply "data about data". Descriptive metadata is contextual data about the model object(s) and documents its characteristics. Metadata is used for the indexing, discovering, and identification. A descriptive metadata record serves several functions: user discovery of an object, access to an object and the management of an object.
Model Pedigree	Model-associated information that describes the model origin, development process, originators and developers, assumptions, expert knowledge, model enhancements, investment costs, versions, change history, etc.

Model Curation: Discussion and Future Research

In this phase of IMCSE research the topic of model curation has been explored through the literature, through technical exchanges, and in discussion with various research stakeholders. The conclusion of the research team is that this is a topic that requires further development in the next phase, engaging members of the systems community. There are opportunities to work with FFRDCs, government agencies and professional societies to mature this area. A number of future opportunities exist, including:

- Define and gain agreement on a model curation lexicon
- Establish a standard model pedigree for systems engineering
- Converge on defined roles and responsibilities for model curation professionals
- Examine alternative architectures, and business case, for centralized vs decentralized curation
- Develop a competency profile for model curation experts

CONCLUSION

IMCSE research seeks to inform and contribute new knowledge, processes, methods and tools to improve the interactivity of humans and models in support of systems decision making. The research is grounded in two assumptions shown in Figure 29. The first is that systems success depends upon effective “human-model teaming” and the second is that specialized leadership and competencies will be required to realize the vision of model-centric engineering.

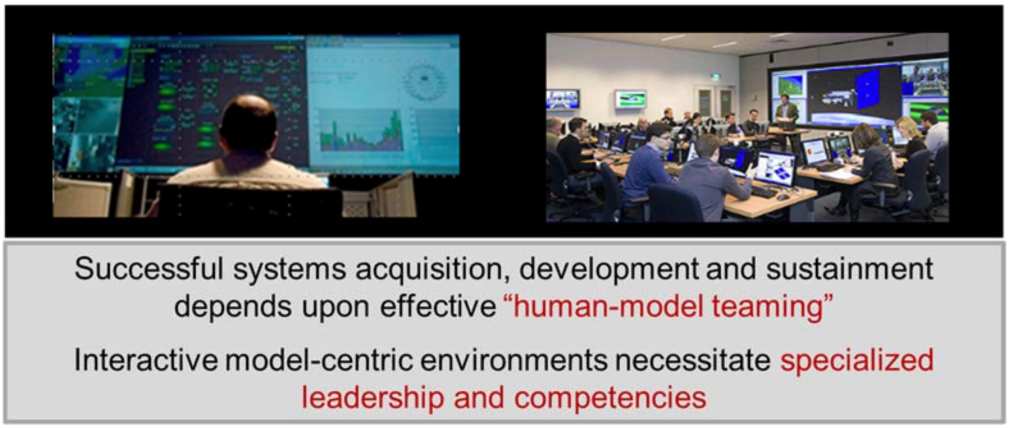


Figure 29 IMCSE Research Assumptions

The phase 4 research program has continued prior phase work, including multi-faceted investigation into human-model interaction. Maturation and further case application of the Interactive Epoch-Era Analysis framework and visualization prototypes was completed. Recommendations on Framing Multi-Stakeholder Tradespace Exploration have been finalized and published. The research team pushed ahead into new areas including an empirical investigation of model-centric decision making and a deeper investigation of model curation as an important capability for model-centric engineering enterprises.

IMCSE research has been presented and discussed with practitioners and sponsors in numerous research meetings and workshops, as well as with other researchers in the systems community. A SERC Talk highlighted various research activities under the project. These activities have raised the awareness of challenges and needs surrounding human-model interactivity, and there is a growing community of interest with the SERC and the larger systems community.

Looking ahead to the next phase of the research program, the team will bring together the prior phase work into an Interactive EEA framework, demonstration prototypes, case-based impact studies, and results of the experiment into a collective set of knowledge artifacts, and will be made available online. The prototype framework with supporting tools will be investigated as a means to support decisions regarding modularity to enable acquisition agility for major defense acquisition programs. The experiment to investigate the impacts of visualization and interaction in a decoupled manner will be completed.

The ongoing empirical study on model-centric decision making will be completed and published in the next phase of the research. Efforts will focus on refinement of the emerging heuristics/design principles, with a goal of developing a validated set of guiding principles for effective human-model interaction. A state-of-the-practice white paper based on the outcomes of the multi-faceted investigation of human-model interaction will be developed. A technical exchange workshop will be convened to develop a set of strategic initiatives based on the research outcomes, with the goal of improving the state-of-the-practice for managing human-model interactivity.

In the next phase of research, the team plans to investigate alternative approaches for transforming organizations to establish leadership and practices for model curation. This will include examining whether different types of organizations need to implement different approaches to model curation. As part of this effort, the goal will be to develop an instrument for organizations to assess their curation capabilities and risks for model-centric environments. A technical exchange workshop to validate outcomes of the model curation research will be convened with interested stakeholders.

Appendix A: List of Publications Resulted

Curry, M. & Ross, A.M. (2016, March). Designing System Value Sustainment using Interactive Epoch Era Analysis: A Case Study for On-orbit Servicing Vehicles. 14th Conference on Systems Engineering Research, Huntsville, AL.

German, E.S. & Rhodes, D.H. (2016, March). Human-Model Interactivity: What Can Be Learned from the Experience of Pilots Transitioning to Glass Cockpit?. 14th Conference on Systems Engineering Research, Huntsville, AL.

Reid, J.B. & Rhodes, D.H. (2016, March). Digital System Models: An Investigation of the Non-technical Challenges and Research Needs. 14th Conference on Systems Engineering Research, Huntsville, AL.

Ross, A.M., Fitzgerald, M.E., & Rhodes, D.H. (2016, March). Interactive Evaluative Model Trading for Resilient Systems Decisions. 14th Conference on Systems Engineering Research, Huntsville, AL.

Reymondet, L., Rhodes, D.H., & Ross, A.M. (2016, April). Considerations for Model Curation in Model-Centric Systems Engineering. 10th Annual IEEE Systems Conference, Orlando, FL.

Fitzgerald, M.E. & Ross, A.M. (2016, July). Recommendations for Framing Multi-Stakeholder Tradespace Exploration. INCOSE International Symposium 2016, Edinburgh, Scotland.

Rhodes, D.H. & Ross, A.M. (2016, July). A Vision for Human-Model Interaction in Interactive Model-Centric Systems Engineering," INCOSE International Symposium 2016, Edinburgh, Scotland.

APPENDIX B: CITED AND RELATED REFERENCES

NSF (2016, Jan 13-14). Research Challenges in Modeling & Simulation for Engineering Complex Systems. National Science Foundation Workshop Report.

Bankes, S. (1993). “Exploratory Modeling for Policy Analysis”, in Operations Research. Vol. 41. No. 3.

Blackburn, M., Bone,M. & Witus, G. (2015, Nov) “Transforming System Engineering through Model-Centric Engineering.” Stevens Institute of Technology, SERC-2015-TR-109.

Blackburn, M. Cloutier, R., Witus, G., Hole, E., & Bone, M. (2015, Jan). “Transforming System Engineering through Model-Centric Engineering.” Stevens Institute of Technology. SERC-2015-TR-044-3.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and Systematic Information Processing within and beyond the Persuasion Context, in *Unintended Thought*. ed. J. S. Uleman & J. A. Bargh. Guilford Press.

Clifford, M.M, Blackburn, M., Verma,D. & Zimmerman, P. (2016, Jul). “Model-Centric Engineering - Insights and Challenges: Primary Takeaways from a Government-Industry Forum.” Stevens Institute of Technology.

Curhan, J. R., Neale, M. A., & Ross, L. (2004). Dynamic valuation: Preference changes in the context of face-to-face negotiation. *J. Exp. Soc. Psychol.*. vol. 40. no. 2. pp. 142–151.

Curry, M. & Ross, A.M. (2015, Mar). "Considerations for an Extended Framework for Interactive Epoch-Era Analysis," 13th Conference on Systems Engineering Research. Hoboken, NJ.

Curry, M. & Ross, A.M. (2016, Mar). "Designing System Value Sustainment using Interactive Epoch Era Analysis: A Case Study for On-orbit Servicing Vehicles". 14th Conference on Systems Engineering Research, Huntsville, AL.

Curry, M. (2017, June). Doctoral Dissertation. Massachusetts Institute of Technology.

Duh, K., Hirao, T., Kimura, A., Ishiguro, K., Iwata, T., & Yeung, C. M. A. (2012, May). “Creating Stories: Social Curation of Twitter Messages”. ICWSM.

Ehrmann, J. R. & Stinson, B. L. (1999). Joint Fact-Finding and the Use of Technical Experts, in *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement*, ed. L. E. Susskind, S. McKearnan, & J. Thomas-Lamar, Sage Publications.

Erikstad, S.O. & Rehn,C.F. (2015) “Handling Uncertainty in Marine Systems Design – State-of-the-Art and Need for Research”. in IMDC 2015.

European Bioinformatics Institute (EMBL-EBI), http://www.ebi.ac.uk, accessed 21 October 2015.

Fisher, R., Ury, W., &Patton, B. (1991*). Getting to Yes: Negotiating an Agreement without Giving In*, Penguin Books, New York NY.

Fitzgerald, M. E. & Ross, A. M. (2014). Controlling for Framing Effects in Multi-Stakeholder Tradespace Exploration*, 12th Conference on Systems Engineering Research*, Redondo Beach CA.

Fitzgerald, M. E. & Ross, A. M. (2015, March). Effects of Enhanced Multi-party Tradespace Visualization on a Two-person Negotiation," *13th Conference on Systems Engineering Research*, Hoboken, NJ.

Garber, M., Sarkani, S., & Mazzuchi, T. A. (2015). Multi-Stakeholder Trade Space Exploration Using Group Decision Making Methodologies, in *25th Annual INCOSE International Symposium*.

Garcia, J.J, Brandt, U.B., & Brett, P.O. (2016, Sep) "Unintentional consequences of the golden era of the Offshore Oil & Gas industry". International Conference on Ships and Offshore Structures.

Garcia, J.J., Pettersen, S.S., Rehn, C.F., & Ebrahimi, A. (2016). 'Handling Commercial, Operational and Technical Uncertainty in Early Stage Offshore Ship Design'. Conference on System of Systems Engineering. Kongsberg, Norway.

Gass, Saul I, & Lambert S Joel. (1980) "Concepts of Model Confidence." *Computers and Operations Research*. 8, No. 4, 341-346.

Gelfand, M. J., Fulmer, A. C., & Severance, L. (2004). The Psychology of Negotiation and Mediation, in *APA Handbook of Industrial and Organized Psychology*, Vol 3, pp. 495–554.

German, E S., & Rhodes, D.H., (2016, Mar). "Human-model interactivity: what can be learned from the experience of pilots with the glass cockpit?" 14th Conference on Systems Engineering Research, Huntsville, AL.

German, E.S. (2017, June) TPP SM Thesis, Massachusetts Institute of Technology.

Goerger S., Madni A., & Eslinger O. (2014, March). "Engineered resilient systems: a DoD perspective," 12th Conf on Sys Eng Research. Redondo Beach, CA.

Golkar, A. & Crawley, E. F. (2014). Framework for Space Systems Architecture under Stakeholder Objectives Ambiguity, *Syst. Eng*., vol. 17, no. 4, pp. 479–502.

Heer, J. & Agrawala, M. (2007). "Design Considerations for Collaborative Visual Analytics," Information Visualization, 7(1):49–62,.

Heer, J. & Shneiderman, B. (2012). "Interactive Dynamics for Visual Analysis," ACM Queue.

Kahneman, D. & Tversky, A. (2000). *Choices, Values, and Frames*. Cambridge University Press.

Kahneman, D. (2011). *Thinking Fast and Slow*, Farrar Straus and Giroux.

Katz-Navon, T. Y. & Goldschmidt, C. (2009, March). Goal orientations in negotiations: the influence of goal orientations on fixed-pie perceptions and bargaining outcomes., *Int. J. Psychol*., vol. 44, no. 1, pp. 60–70.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*, University of Chicago Press.

Kumar, Ranjit. (2011) *Research Methodology: a step-by-step guide for beginners*. (3rd ed.). London: SAGE.

La Tour, P.A. & Hastings, D. E. (2015). “Combining System Dynamics with Tradespace Exploration to Explore Future Space Architectures”, IAC.

Lee, John D. & Katrina A See. (2004). “Trust in Automation: Designing for Appropriate Reliance.” Humans Factors. vol. 46, no. 1, 50-80.

Levin, I., Schneider, S., & Gaeth, G. (1998, November). All Frames Are Not Created Equal: A Typology and Critical Analysis of Framing Effects, *Organ. Behav. Hum. Decis. Process*., vol. 76, no. 2, pp. 149–188.

Liu Z, Jiang B, Heer J. (2013). “Immens: Real-Time Visual Querying of Big Data”, Eurographics Conf on Vis (EuroVis).

Mathieu, J. Melhuish, J. James, J., Mahoney, P., Boiney, L., White, B. (2007). “Multi-Scale Modeling of the Air and Space Operations Center”, The MITRE Corporation.

McBurney, P. (2011). “What are Models for?,” in: M. Cossentino, K. Tuyls & G. Weiss (Editors): Post-Proceedings of the Ninth European Workshop on Multi-Agent Systems, Lecture Notes in Computer Science, volume 7541. Berlin, Germany: Springer, pages 175- 188,

Ozawa, C. P. (1991). *Recasting Science: Consensual Procedures in Public Policy Making*, Westview Press.

Pace, Dale K. (2004). “Modeling and Simulation Verification and Validation Challenges.” Johns Hopkins APL Technical Digest.

Parasuraman, Raja, & Victor Riley. (1997). “Humans and Automation: Use, Misuse, Disuse, Abuse.” Human Factors. 1997: vol. 39, no. 2, 230-253.

Petty, R. & Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*, Springer.

Raiffa, H. (2002). *Negotiation Analysis: The Science and Art of Collaborative Decision Making*, Belknap Press of Harvard University Press.

Rawls, J. (1971). *A Theory of Justice*, Harvard University Press, Cambridge MA.

Rehn, C.F., Pettersen, S.S., Garcia, J.J., Erikstad, S.O., Brett, P.O., Asbjørnslett, B.E., Ross, A.M., & Rhodes, D.H. (2016) Ill-Structured Commercial Ship Design Problems: The Responsive System Comparison Framework on an Offshore Vessel Case, Working Paper. MIT. Cambridge, MA.

Reymondet, L., Rhodes, D.H., & Ross, A.M. (2016, April). Considerations for Model Curation in Model-Centric Systems Engineering. 10th Annual IEEE Systems Conference, Orlando, FL.

Rhodes D.H. & Ross, A.M. (2015, Feb 28). *Interactive Model-Centric Systems Engineering Technical Report,* SERC-2014-TR 048-2

Rhodes, D.H. & Ross, A.M. (2015, Feb 28). *Interactive Model-Centric Systems Engineering Pathfinder Workshop Report*, MIT, Cambridge, MA.

Rhodes, Donna H. & Ross, A.M. (2016). "A Vision for Human-Model Interaction in Interactive Model-Centric Systems Engineering." presented at the INCOSE International Symposium 2016, Edinburgh, Scotland.

Rhodes, Donna H. & Ross, A.M. (2015, March). "Interactive Model-Centric Systems Engineering (IMCSE) Phase 3." Stevens Institute of Technology, SERC-2015-TR-043-1.

Ricci, N., Schaffner, M. A., Ross, A. M., Rhodes, D. H., & Fitzgerald, M. E. (2014, March). Exploring Stakeholder Value Models Via Interactive Visualization, *12th Conference on Systems Engineering Research*, Redondo Beach, CA.

Roberts J. (2007). "State of the Art: Coordinated & Multiple Views in Exploratory Visualization," 5th Int'l Conf on Coordinated and Multiple Views in Exploratory Vis, Washington, DC.

Ross A.M. & Rhodes D.H. (2008, June). "Using Natural Value-Centric Time Scales for Conceptualizing System Timelines through Epoch-Era Analysis". INCOSE Int'l Symp 2008. Utrecht, the Netherlands,.

Ross, A. M. & Hastings, D. E. (2005, July). The Tradespace Exploration Paradigm, in *INCOSE International Symposium 2005*, Rochester, NY.

Ross, A. M., McManus, H. L., Rhodes, D. H., & Hastings, D. E. (2010a, September). Revisiting the Tradespace Exploration Paradigm: Structuring the Exploration Process, AIAA Space 2010, Anaheim, CA.

Ross, A. M., McManus, H. L., Rhodes, D. H., & Hastings, D. E. (2010b, September). A Role for Interactive Tradespace Exploration in Multi-Stakeholder Negotiations, AIAA Space 2010, Anaheim, CA.

Ross, A. M., Rhodes, D. H., & Fitzgerald, M. E. (2015, March). Interactive Value Model Trading for Resilient Systems Decisions, *13th Conference on Systems Engineering Research*, Hoboken, NJ.

Rouse, W.B. (2015).“Modeling and Visualization of Complex Systems and Enterprises,” Exploration of Physical, Human, Economic and Social Phenomena.

Rusbridge, C., Buneman, P., Burnhill, P., Giaretta, D., Ross, S., Lyon, L., & Atkinson, M. (2005). "The Digital Curation Centre: A Vision for Digital Curation", IEEE Intl Symposium on Mass Storage Systems and Technology.

Scherr M. (2009). “Multiple and Coordinated Views in Information Visualization”, Media Informatics Advanced Seminar on Info Vis 2008/2009.

Schon, D. & Rein, M. (1994). *Frame Reflection: Toward the Resolution of Intractable Policy Controversies*, Basic Books.

Scott, M. J. & Antonsson, E. K. (2000). Arrow’s Theorem and Engineering Design Decision Making, *Res. Eng. Des*., vol. 11, no. 4, pp. 218–228.

Shanteau, J. (1992). Competence in experts: The role of task characteristics, *Organ. Behav. Hum. Decis. Process.*, vol. 53, no. 2.

Simon, H. A. (1957). *Models of Man*, Wiley.

Tversky, A. & Kahneman, D. (1974, September). Judgment under Uncertainty: Heuristics and Biases, *Science*, vol. 185, no. 4157, pp. 1124–31.

Tversky, A. & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice, *Science*, vol. 211, no. 4481, pp. 453–458.

West, Timothy D, & Art Pyster. (2015). “Untangling the Digital Thread: The Challenge and Promise of Model-Based Engineering in Defense Acquisition.” INSIGHT 2015; vol. 18, 45-55.

Wilson, T. D. & Schooler, J. W. (1991). Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions, *J. Pers. Soc. Psychol*., vol. 60, no. 2, pp. 1–23.

Zacharias, G., J. MacMillan, J., Van Hemel, S.B. (2008). *Behavioral Modeling and Simulation*, The National Academies Press.

Zimmerman, Philomena. (2015). “MBSE in the Department of Defense.” Office of the Deputy Assistant Secretary of Defense for Systems Engineering, US Department of Defense.

Zulkepli, J., Eldabi, T., Mustafee, N. (2012). “Hybrid Simulation for Modelling Large Systems: An Example of Integrated Care Model”. Proceedings of the 2012 Winter Simulation Conference.